UNIVERSITY *of* TASMANIA

Appropriate Use of Cardiac Imaging in Australia

By

Ricardo Andrés Fonseca Díaz

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (Medical Research)

Menzies Institute for Medical Research

University of Tasmania

Hobart, Australia

March 2017

Supervision

This thesis was supervised by

Primary Supervisor

Professor Thomas H. Marwick, MD, PhD

Menzies Institute for Medical Research

University of Tasmania

Hobart, Australia

And

Co-supervisor

Dr Kazuaki Negishi, MD, PhD

Menzies Institute for Medical Research

University of Tasmania

Hobart, Australia

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and**,** to the best of my knowledge and belief**,** no material previously published or written by another person**,** except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed: _____

3rd March 2017

Date: _____

Statement of Authority of Access

The publishers of the papers included in Chapters 1, 3, 4, 5, 6, 8, and 9 hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining non-published content of this thesis may be made available for loan and limited copying and communication in accordance with the *Copyright Act 1968*.

Signed: _____

3rd March 2017

Date: _____

Statement of Ethical Conduct

All research associated with this thesis abides by the International and Australian codes on human and animal experimentation, and full ethical approval from the relevant institutions was obtained for all studies outlined in this thesis. All individual participants provided written informed consent for involvement in the respective research studies.

Signed: _____

3rd March 2017

Date: _____

Statement of Sponsorship

The research reported in this thesis was supported by an Australian Government Research Training Program Scholarship and The Farrell Foundation Scholarship in Cardiovascular Health and Diseases Research.

Signed: _____

20th July 2017

Date: _____

Statement of Authorship

The following papers are incorporated into chapters of this thesis and were either published or submitted for publication in peer reviewed scientific journals during the course of this PhD candidature.

Author details and their roles:

Paper 1: Located in Chapter 1.

Fonseca R, Marwick TH. How I do it: judging appropriateness for TTE and TEE. Cardiovasc Ultrasoun. 2014;12(1):22.

Author 2 contributed with the idea. The candidate (Author 1) was the first author and with Author 2 designed the manuscript. Author 1 was responsible for data collection and data analysis under supervision of Author 2. Author 1 prepared the manuscript draft. Author 2 edited and approved the final manuscript. All the authors had access to all the data including statistical reports and tables.

Paper 2: located I Chapter 1.

Fonseca R, Marwick TH. Appropriateness and outcomes: is it time to adopt appropriate use criteria outside of North America? Heart. 2014;100(5):357-8.

Author 2 contributed with the idea. The candidate (Author 1) was the first author and with Author 2 designed the manuscript. Author 1 was responsible for data collection and data analysis under supervision of Author 2. Author 1 prepared the manuscript draft. Author 2 edited and approved the final manuscript All the authors had access to all the data.

Paper 3: Located in Chapter 3.

Fonseca R, Otahal P, Wiggins N and Marwick TH. Growth and geographical variation in the use of cardiac imaging in Australia. Intern Med J. 2015;45:1115-27.

Author 4 contributed with the idea. The candidate (Author 1) was the first author and with Author 4 designed the manuscript. Author 1 and Author 3 were responsible for data collection. Author 1 and Author 2 were responsible for statistical analysis under supervision of Author 4.

Statement of Authorship

Author 1 was responsible for geo-mapping. Author 1 prepared the manuscript draft. Author 4 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including statistical reports and tables included in the study.

Paper 4: Located in Chapter 4.

Fonseca R, Negishi K and Marwick TH. What is the evidence status of Appropriate Use Criteria (AUC)? Insight from a matching exercise with the guidelines for echocardiography. Intern Med J. 2015;45:864-9.

Author 3 contributed with the idea. The candidate (Author 1) was the first author and with Author 3 designed the manuscript. Author 1 was responsible for data collection and data analysis. Author 2 and Author 3 supervised data analysis. Author 1 prepared the manuscript draft. Author 3 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including tables in the study.

Paper 5: Located in Chapter 5.

Fonseca R, Negishi K, Otahal P and Marwick TH. Temporal changes in appropriateness of cardiac imaging. J Am Coll Cardiol. 2015;65:763-73.

Author 4 contributed with the idea. The candidate (Author 1) was the first author and with Author 4 designed the manuscript. Author 1 was responsible for data collection and data analysis. Author 2 and Author 4 supervised data collection. Author 2 and Author 3 supervised data analysis. Author 1 prepared the manuscript draft. Author 4 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including tables in the study.

Paper 6: Located in Chapter 6.

Fonseca R, Otahal P, Galligan J, Neilson S, Huynh Q, Saito M, Negishi K and Marwick TH. Association of Survival Time with Transthoracic Echocardiography in Stable Patients with Heart Failure: Is Routine Follow-Up Ever Appropriate? *International Journal of Cardiology*. 2016.

Statement of Authorship

Author 8 contributed with the idea. The candidate (Author 1) was the first author and with Author 8 designed the manuscript. Authors 1,3,4,5 and 6 were responsible for data collection. Author 1 was responsible of data analysis. Author 7 and Author 8 supervised data collection. Author 2 and Author 7 supervised data analysis. Author 1 prepared the manuscript draft. Author 8 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including tables in the study.

Paper 7: Located in Chapter 8.

Fonseca R, Pathan F and Marwick TH. Development and validation of a screening tool for the identification of inappropriate transthoracic echocardiograms. BMJ Open. 2016;6:e012702.

Author 3 contributed with the idea. The candidate (Author 1) and Author 2 were joint first authors. Author 1 and Author 4 designed the manuscript. Author 1 and Author 2 were responsible for data collection and data analysis. Author 3 supervised data collection and data analysis. Author 1 and Author 2 prepared the manuscript draft. Author 3 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including tables in the study.

Paper 8: Located in Chapter 9.

Pathan F, **Fonseca R** and Marwick TH. Usefulness of Hand-Held Ultrasonography as a Gatekeeper to Standard Echocardiography for "Rarely Appropriate" Echocardiography Requests. *Am J Cardiol*. 2016;118:1588-1592.

Author 3 contributed with the idea. The candidate (Author 2) and Author 1 were joint first authors. Author 1 and Author 3 designed the manuscript. Author 1 and Author 2 were responsible for data collection and data analysis. Author 1 was responsible of scans. Author 3 supervised data collection and data analysis. Author 1 and Author 2 prepared the manuscript draft. Author 3 edited the final manuscript. All the authors approved the final manuscript. All the authors had access to all the data including tables in the study.

Signed: _____

Ricardo Andrés Fonseca Díaz (PhD Candidate)

3rd March 2017

Date: _____

Signed: _____

Professor Thomas H. Marwick (Primary supervisor)

3rd March 2017

Date: _____

Table of contents

Table of contents

Table of contents

Table of contents

List of tables

List of tables

List of tables

List of figures

List of figures

List of figures

Acknowledgments

Firstly, I would like to express my most honest appreciation to my primary supervisor Professor Thomas H. Marwick, for accepting me as a PhD candidate; for the continuous support during my research; for his persistence, passion, and immense knowledge. His direction assisted me throughout the investigation and writing of this work. He has helped me to continuously improve myself as a researcher, to think better, to work harder, and to desire higher. I could not have been luckier to have a better primary supervisor than him.

Besides my mentor, I want to thank my co-supervisor Dr Kazuaki Negishi for his contribution; to my workmates Ying, Faraz, Hilda, Quan, Tomoko, and former workmates Andrea, Makoto, Leah, for the fun and support. My experience in this Institute was undoubtedly better because of you. And to Lisa, for her organisational skills, which I would like to learn!

To Mr Petr Otahal, for his invaluable time, patience, guidance, and statistical teaching. I would not have been a success without your help. Thank you for always stopping what you were doing when I needed your support. To Dr Kim Jose, for her priceless time helping me to understand the 'unknown' world of qualitative research. And to Professor James Sharman, for the advice received about how to handle the challenges that life as a researcher brings these days.

To Larissa Bartlett, for her vital contribution to finding me a sponsor. To the Farrell Foundation and especially to Ms Julia Farrell, for agreeing to be my sponsor. I have no words to express my most sincere gratitude to you. I wish you the best and hope you can continue helping people who do not have enough resources to do research for this city, this state, and this amazing country.

To Edwin for convincing me to come to Australia. Thank you for the unconditional support received during all those years. To my "Australian family": Pip, Jill, Jo and Dave; Pat and Gordon; Elisa and John, Diana, Susie, Shelly, Rachel. Thank you for reassuring me and for believing in my capabilities.

To Ben, for being always there when you, with no hesitation, were ready to cheer me and to show me how to be passionate, loyal, and persevering, even in the most difficult times. I have no words to express my gratitude.

To other staff at the Menzies Institute for Medical Research, especially to Griffin, Amanda, Miranda, Alison, and Mark, for making me feel valuable, welcomed, and for giving me the best environment to achieve my purposes.

Acknowledgments

And last but not the least, my love and deepest gratitude to my parents, Alvaro and Olga; my sister and my brother, Silvia and Camilo, for supporting me spiritually throughout this PhD, to my beautiful nieces Luciana and Maria Alejandra, to my aunts and uncles, cousins, and friends in Colombia. Ustedes no saben cuánta falta me hacen, cuánto los extraño, y cuánto quisiera tenerlos siempre a mi lado.

Abstract

Health systems around the world are concerned by the same topic: How their national health budgets can be spent efficiently and economically[1]. Some related factors, such as the ageing population, an increase in sophistication and costs of treatments and an increase in the demand for imaging services, have all contributed to the rise in healthcare spending, placing further burdens on limited resources. The healthcare system of the United States of America has found this particularly challenging, where the use of cardiovascular imaging had grown persistently for the previous two decades[1]. This increase, associated with marked geographical variation of imaging utilisation, the spread of imaging equipment, repetition of use due to poor quality tests, and defensive medicine, have led to the concern about the value of cardiac imaging[1, 2]. Part of the response in the United States of America to the disconnection between the use in imaging and its value has been the development of the Appropriate Use Criteria (AUC) for Cardiac imaging[1]. The first criteria were introduced in 2005, with subsequent versions being published to include different cardiac imaging techniques, aiming to improve utilisation of imaging technologies in an efficient way, contributing to improvement in patient care, doctors' decision making, standardisation of medical practice and control of health expenditure. The concept of appropriate use of cardiovascular imaging has remarkably influenced the relationship between clinicians, health policymakers and insurance companies to the point that from January 2017, all doctors and hospitals must certify their use and adherence to the criteria in order to receive reimbursement under the United States Medicare schedule[3].

Despite the utilisation of the AUC for the past 12 years, several problems are apparent. Issues related to the scientific evidence of the criteria; the uncertainty of the impact of the AUC on clinicians' requesting behaviour, and on health outcomes, question the widespread use of them to achieve appropriate use of cardiovascular imaging in the United States and other countries like Australia.

Due to an increased interest in the utilisation of these criteria in the Australian practice, this thesis which comprises quantitative and qualitative research methods, investigates the use of cardiac imaging techniques in Australia and if the American Appropriate Use Criteria are suitable to use in the Australian health care system.

The work in this thesis is divided into three parts. The first part (Chapter 3), aims to determine if the use of cardiac imaging in Australia is appropriate based on the analyses of growth and geographical variation of imaging used in different regions of the country. Data from Medicare

Abstract

Australia Statistics, Australian Health Survey, and Health Workforce were analysed to define the growth and regional variation in the use of imaging[4]. The increase of imaging was determined using rate of tests per 100,000 people in each geographical area (Medicare locals) and Spearman correlations, negative binomial regression and zero-inflated negative binomial regressions, were used to investigate univariable and multivariable associations between age-weighted testing and characteristics of each Medicare local[4]. The results identified the ongoing growth of cardiac imaging use per 100,000 people in Australia, more pronounced in echocardiography modalities (transthoracic, trans-oesophageal, and stress echocardiography). Stress echocardiography was the technique with the highest increase since 2002 (423%), and there was a substantial regional variation in the use of all echocardiography modalities[4]. Greater regional echocardiography use was associated with females, proximity to major cities, higher socioeconomic status, and local concentration of physicians, but not with disease burden or deaths due to cardiovascular causes. The medical workforce appears to be the strongest independent driver of echocardiography use[4].

The next section (Chapters 4 to 7) aims to determine whether the Appropriate Use Criteria are suitable for use in Australia to achieve appropriate use of imaging. Several studies were undertaken to understand this. First, we related similarities and relationships between cardiac guidelines and Appropriate Use Criteria; second, we assessed the impact of the criteria on clinicians' requesting behaviour. Third, we evaluated the impact on health outcomes, and finally, factors that lead doctors to order "inappropriate". echocardiograms.

For examining the similarities of Appropriate Use Criteria with guidelines (Chapter 4), the criteria and the published Cardiology guidelines were matched to identify concordance between both. Concordance was determined when "appropriate" items in the criteria had recommendation class I or IIa in the guidelines; "inappropriate" scenarios had recommendation class III, and "uncertain" items had recommendation class IIb[5]. The results identified that 91% of the criteria had a counterpart on the guidelines for the use of echocardiography[5]. However, only 82% of them were concordant, and the rest were inconsistent. These results suggest that the potential incorporation of an AUC process into Australian practice might face challenges to achieve appropriate use and to improve quality of care[5].

Chapter 5 was a systematic review of published manuscripts related to the AUC use in cardiovascular imaging to study the impact of the appropriateness criteria on clinicians' ordering behaviour. The aim was to determine the proportion of "appropriate" and

Abstract

"inappropriate" testing over time using meta-regressions. More than 5,200 manuscripts were found in online databases and were analysed. The results showed that there was a positive association with the percentage of appropriate tests for Transthoracic Echocardiography (TTE), as well as a negative association of the ratio of uncertain tests over time, but inappropriate testing remained the same, that is no decrease in the proportion of inappropriate tests over time. These results suggest no impact on clinicians' ordering behaviour[6].

Chapter 6 aimed to determine the effect of the definition of appropriateness on survival time in stable heart failure patients. The analysis was made using survival analysis with time-dependent variables for the combined endpoint of Heart Failure (HF) readmission and death, and a separate analysis for HF readmission, with death as a competing risk[7]. The results showed no differences in the event-free time for combined outcomes; HF readmission was not associated with routine follow-up TTE timing; there were no differences in the cumulative incidence of death between groups[7].

For chapter 7, a qualitative analysis was performed using semi-structured face-to-face interviews with open-ended questions to analyse the factors that lead doctors to order echocardiograms and the relationship with inappropriate testing. Seventeen physicians (cardiologists and non-cardiologists) participated in the study. Personal factors such as lack of expertise, limited experience and inability to manage uncertainty were the most important factors that impacted the decisions of doctors leading to the ordering of inappropriate tests. Other factors included accessibility of tests and adherence to protocols. These results suggest a mismatch between the clinical reasoning of physicians and the AUC for echocardiography.

Due to the results of the previous chapters indicating that the AUC are not the ideal tools to achieve appropriate use, to improve health outcomes, and to change doctors' ordering behaviour, a third and final section evaluated a proposed methodology to determine possible inappropriate testing and the way to give answers to doctors without the need for performing a formal test. This third section includes chapters 8 and 9.

For Chapter 8, general characteristics of inappropriate TTE were determined and summarised in a questionnaire of four binary questions most commonly associated with inappropriate tests. These questions are related to the absence or change of cardiovascular symptoms or signs, routine surveillance as the purpose of the exam, echocardiograms in the previous twelve months, and if endocarditis with no murmur or positive blood cultures was the reason for the scan[8]. These questions were applied to two different validation groups. Analyses of specificity,

sensitivity, and predictive values were performed to determine the accuracy of the prediction of inappropriate requests. Two or more affirmative questions had a high sensitivity and specificity to determine inappropriate requests according to the criteria. The time to answer the questions was limited to up to 2 minutes when it could not be answered directly from the echo request. In only 19% of the requests was there a need to the review the digital record to respond to the questions. Only 18% of the total amount of requests to the echo laboratory, warranted a comparison with the Appropriate Use Criteria[8].

Two approaches to inappropriate requests (standard TTE versus hand-held echocardiography) were compared in a case-control study (Chapter 9). Patients were followed for a period of 6 months and outcomes such as time until the scan, repeat echocardiography, the cost of each strategy, new findings, and impact on management were determined[9]. The results of this study showed no differences in the finding of new pathologies or change in management between hand-held echocardiography and standard TTE. However, people in the portable echo group had less time to scan if they were inpatients and the average cost of the approach was lower than the standard echocardiography group ($145 AUD vs $241 AUD)[9].

Statistical analyses in all chapters were performed using R software; geo-mapping for Chapter 2 was done using QGIS 2.4.0.; hand-held ultrasonography was performed using the GE Vscan V 1.2 portable device and standard TTE was performed GE Vivid 9, Philips IE33 and Acuson SC2000.

Conclusions

The results of the studies contained within this thesis suggest there is a need for appropriate use of cardiac imaging in Australia because there is a vast difference of use in Medicare locals and no association of cardiac imaging use with the burden of cardiovascular disease. However, the use of the American AUC may face critical challenges for use in Australia because of a number of factors, these being: the discordance of the criteria with the cardiology guidelines, the lack of evidence of impact of the Appropriate Use Criteria on physicians' ordering behaviour, the absence of improvement of survival or readmission in patients with stable heart failure, and the mismatch between the clinical reasoning process of doctors and the criteria[4-7].

Those issues aside, our studies found that the use of a simple questionnaire at the point-of-service assists in determining a high proportion of inappropriate tests in an unbiased way with minimal interruption of workflow. The use of hand-held echocardiography appears to be a safe and cost-saving strategy that can be used to give the answers to doctors without the need for

performing standard echocardiograms. The combination of these, enable point-of-service staff to determine appropriateness and provide a viable alternative to costly tests, helping to reduce the economic burden while still providing doctors with access to information in support of their patients' care[8, 9].

Chapter 1. Introduction

Part of the research contained in this chapter has been published as[1, 2]:

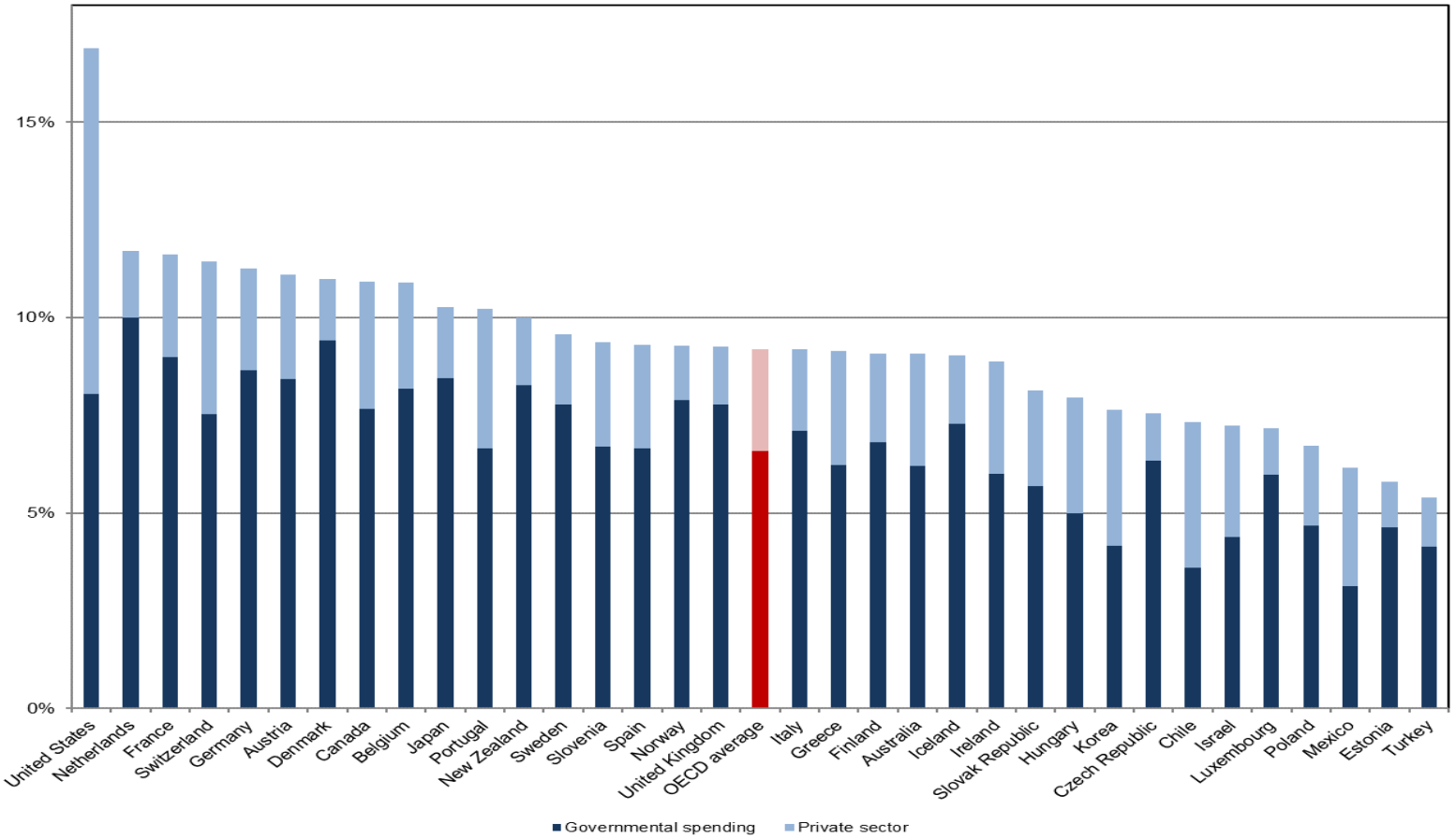
- Fonseca R, Marwick TH. How I do it: judging appropriateness for TTE and TEE. *Cardiovasc Ultrasoun*. 2014;12(1):22.
- Fonseca R, Marwick TH. Appropriateness and outcomes: is it time to adopt appropriate use criteria outside of North America? *Heart*. 2014;100(5):357-8.

One of the most challenging tasks during my years of experience as a physician and as a specialist of healthcare administration has been to balance the intricate relationship between the practice of medicine *per se* and the economic concept of it. Unfortunately, medical practice is no longer simply a service to our patients with a strong code of ethics; we are also finding ourselves incresingly bound by finite resources and laws of economy that drive decisions of care in our practices and healthcare systems. Our actions as medical practitioners are readily measured by analysing health care expenditure against health outcomes. It is difficult to identify the proportion of practitioners acutely aware of the breadth of economic implications their work causes to society. And it is here, with the unrecognised concepts of appropriate use and value of healthcare, that the battle between governments, managers, economists, and policymakers vs. health professionals, has its origins.

The rise of healthcare expenditure: the cornerstone

Health care spending has been increasing worldwide regardless of wealth, population characteristics, or region [10-12]. This trend is marked in developed countries since the 1990's [10, 13, 14], the United States of America being the country with the highest health expenditure in the world [11] (Figure 1- 1).

Figure 1- 1 Health spending as a share of GDP, 2012 or latest year

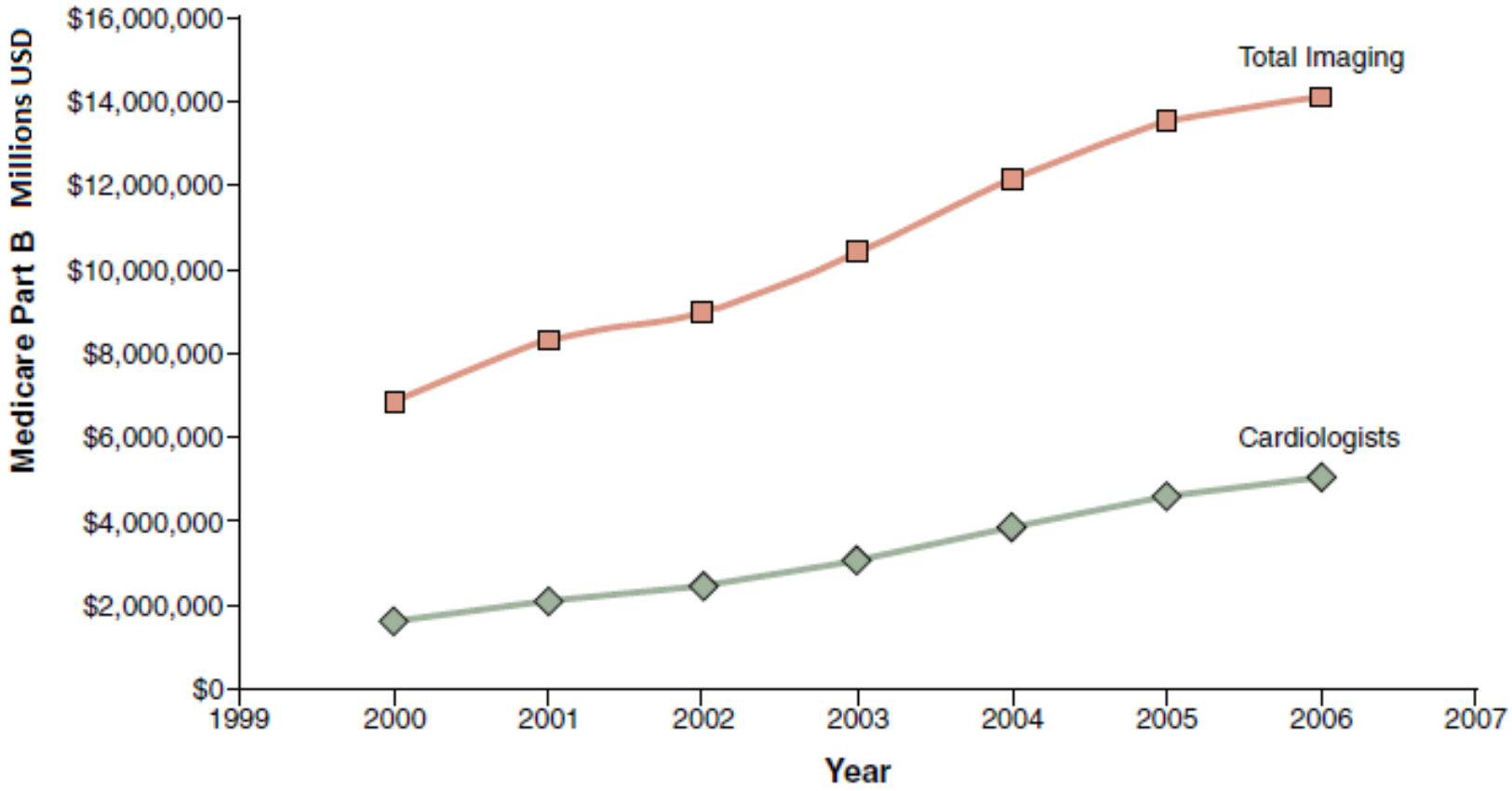


Reproduced from OECD. Fiscal Sustainability of Health Systems: OECD Publishing[11]

Although the medical-care industry's problems have been analysed for decades[15], a particular situation occurred that "set the alarm bells ringing" in recent years. This situation was the increase of imaging use and its significant impact on health budgets, raising the question as to how the imaging component of the health budget can be spent wisely[1].

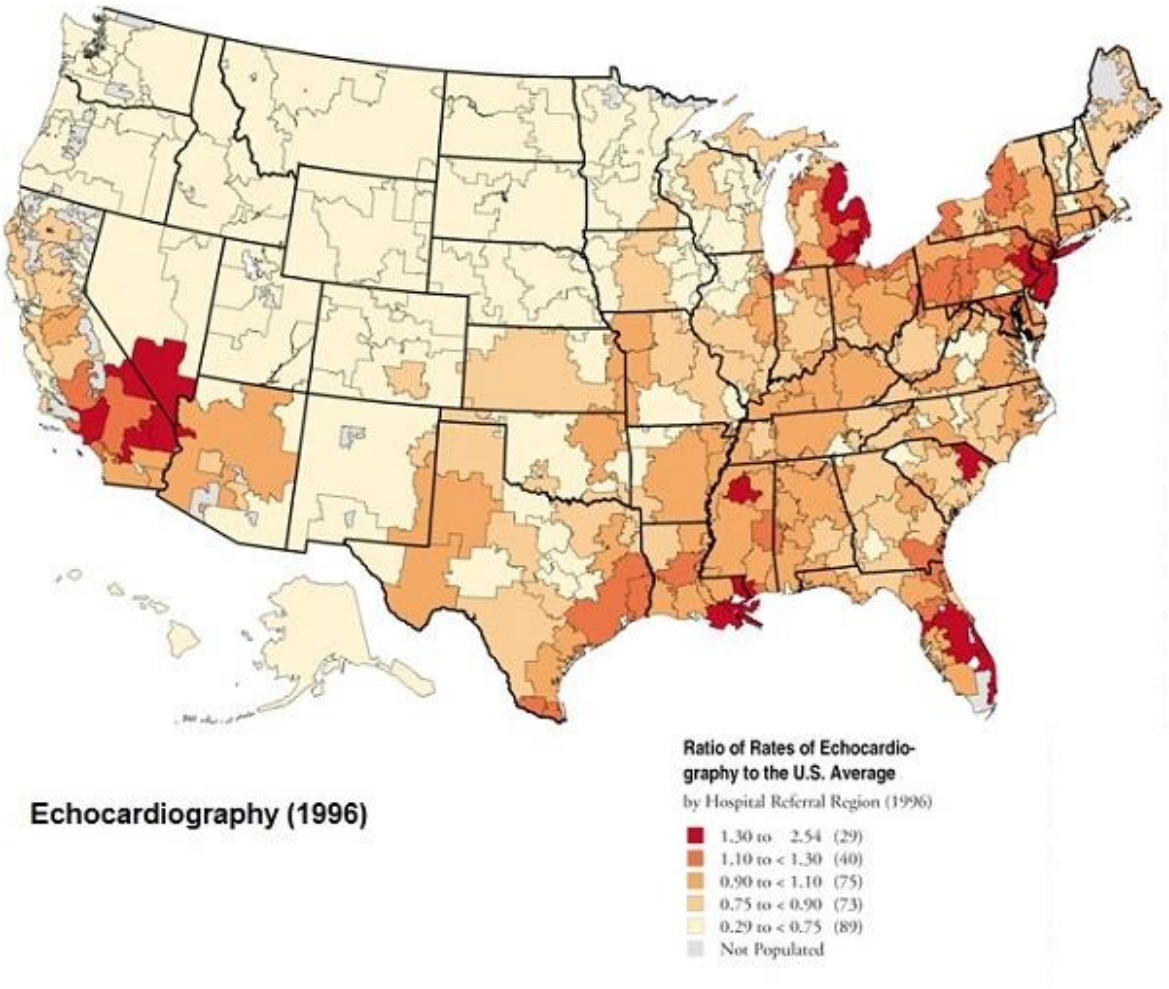
There is perhaps no space where this is more challenging than cardiovascular imaging, the growth of which has exceeded the overall increase of medical costs over the last decades, which has been more pronounced in the United States of America (USA)[1, 11, 16]. The contribution of imaging to the medical budget started to be highlighted in the United States more than 20 years ago. At this time, the Medicare Payment Advisory Commission (MedPAC) showed a 10%/year increase in spending for cardiac imaging between 1999 and 2002, when the average growth per year of all services was 5.2%[2, 17]. This trend continued throughout the following decade: the imaging payments to Cardiologists in 2000 were US$1.6 billion, increasing to US$5.1 billion in 2006 [2, 18] (Figure 1- 2).

Figure 1- 2 Medicare reimbursements to cardiologists in the US (Reproduced from Shaw, et al.[18]).



Contributors to this growth included the rapid proliferation of imaging machines, limited experience with new imaging modalities among non-specialists, automated referral pathways and self-referrals [2, 19], poor quality of imaging (requiring repetition of tests) and defensive medicine [2, 20, 21]. In addition to these previous contributors, the differences in the use of imaging amongst regions supported the contention that the selection of imaging tests was discretionary rather than disease-related [22-24] (Figure 1-3), rising concerns about the poor value of imaging tests.

Figure 1- 3 Differences in the use of echocardiography in the US in 1996. Regional variations by hospital referral region expressed as a ratio to the US average. From Wennberg, et al. The Dartmouth Atlas of Cardiovascular Health Care. P65. 1999 [24].



Strategies to content use of cardiac imaging

Several strategies have been implemented in an attempt to control the use of cardiac imaging and reduce spending in the United States. These approaches include (but not limited to):

- Changes in reimbursement models, which aims to reduce refund costs of cardiac imaging to discourage the overuse of tests[18];
- Use of Radiology Benefit Manager Companies by health insurers, which adopt pre-authorisation processes seeking to deny payment of services if no pre-authorisation has been granted[18];
- Adoption of the Appropriate Use Criteria [18, 25].

Development and application of the Appropriate Use Criteria

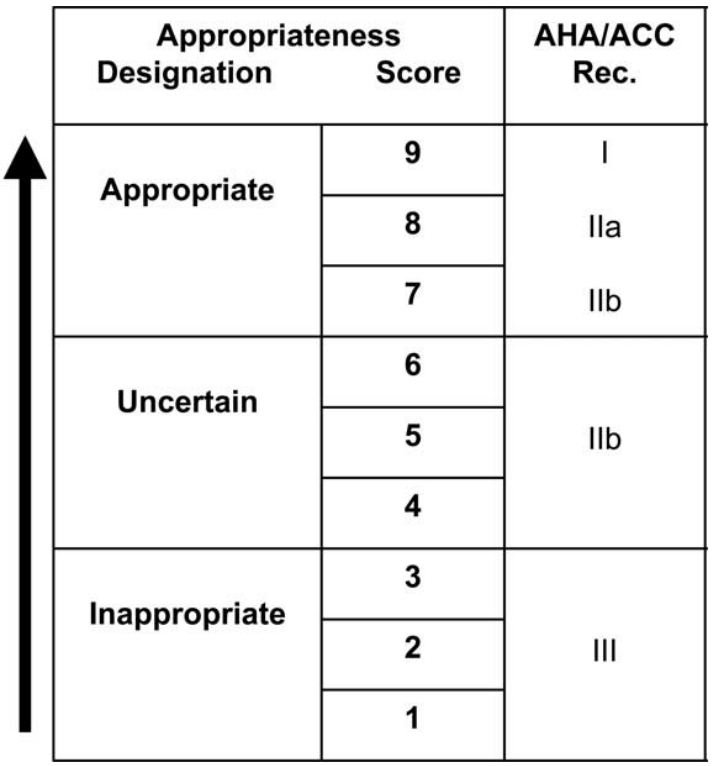
Part of the response to the disconnect between the growth in imaging and its value in North America has been the development of the Appropriate Use Criteria (AUC)[1]. The criteria were designed to improve physicians' decision making, to enhance patient's care, and to control the use of imaging and therefore, health expenditure[1, 26].

Although the initial application of Appropriate Use Criteria was not directed towards cardiovascular imaging but towards medical procedures[27], the American College of Cardiology Foundation (ACCF), along with other medical associations, formed the Appropriateness Criteria Working Group (now called ACCF AUC Task Force) [2, 28]. The ACCF AUC Task Force used a modified Delphi process (based on the Delphi method employed by the RAND/UCLA appropriate use criteria methodology) [2, 27, 29] to select the appropriateness criteria.

The original Delphi process (in respect to healthcare) employed by the RAND Corporation ("Research and Development") and the University of California, Los Angeles (UCLA), was developed in the 1980s to address the problem of overuse and underuse of health procedures[27]. The aim of the method was to "*combine the best available scientific evidence with the collective judgment of experts to yield a statement regarding the appropriateness of performing a procedure at the level of patient-specific symptoms, medical history, and test results*" [27]. The original methodology involves a list of questions being compiled and sent to a panel of experts who answer anonymously. Their responses are collated and summarised to reflect the "group response". This summary is sent back to the members of the panel as statistical analysis, to protect respondent anonymity. The process is repeated, over some iterations, to formulate an "expert" collective agreement[27].

The modified Delphi process used to develop the Appropriate Use Criteria for different cardiac imaging technologies is an evaluation of a list of medical indications (in which a cardiac imaging test could be performed) by a panel of experts (comprising 9 to 15 individuals who are physicians, health professionals, and policy makers)[30]. A review committee compiles a list of clinical scenarios after assessing the scientific evidence available [30]. Individually, the expert panel members analyse the risks, benefits and recommendations of evidence for performing the imaging test in each medical situation. They then indicate their response by providing a score between 1 to 9 according to the appropriateness of the test. In their review of the medical cases, the panel members also reference the Cardiac Guidelines to assist in determining their grading for appropriateness[30] (Figure 1- 4).

Figure 1- 4 Guides for panel reviewers to consider (Reproduced from Patel, et al.[30])

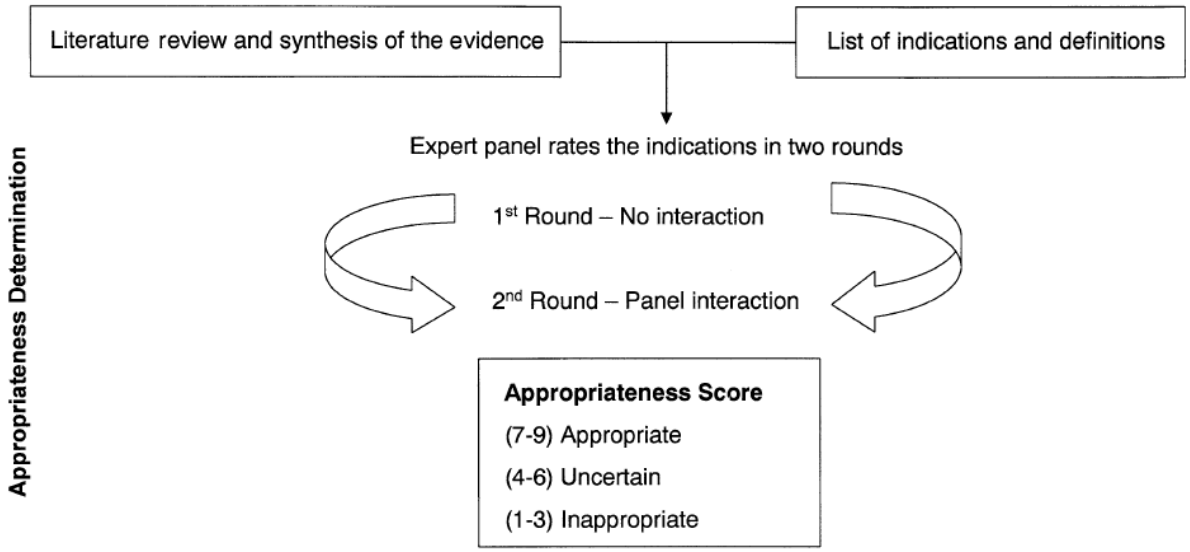


An assessor gives a score between 7 to 9 to what they deem to be an "appropriate" test for that particular indication. In this case, the test is considered mostly acceptable and is an adequate approach for the clinical scenario or indication. When the score is between 1 to 3, the test would

not be considered broadly "*acceptable"* and therefore "*not a reasonable approach"* for the situation ("Inappropriate")[26, 30]. Finally, scores between 4 to 6, identify those "uncertain" or "may be appropriate" tests for that particular situation. In this case, the test may be acceptable and accordingly may be considered a reasonable approach for the situation. However, an "uncertain" test may be the result of ambiguity which implies that more research or information is necessary to determine the status as appropriate or inappropriate[30].

After the first ranking round as described above (done privately by each panel member without interactions with other individuals of the group), a second round is done to reach consensus among the members of the committee. This meeting is conducted collectively, to allow discussion and debate on the various scenarios before a final ranking is submitted. The averages of the ranks are split between 7 to 9, 4 to 6 and 1 to 3 following the second round. The indications, respective to the splits noted above, are defined as "appropriate", "uncertain" or "inappropriate" (the latter two now being referenced as "may be appropriate" and "rarely appropriate" by the new terminology)[28, 30]. (Figure 1- 5).

Figure 1- 5 Process to determine appropriateness based on RAND/UCLA method (Reproduced from Patel et al. [30])



The concept of appropriateness defined by the RAND/UCLA methodology in the 1980's was the cornerstone for developing the first attempt at appropriate use criteria (AUC)[2]. The initial concept suggested that "*an appropriate procedure is one in which 'the expected health benefit (e.g., increased life expectancy) exceeds the expected negative consequences (e.g., mortality, morbidity, anxiety, pain, time lost from work) by a sufficiently wide margin that the procedure is worth doing, regardless of cost*"[2, 27, 29]. However, no management decisions in medical practice are exempt from having to address an abstract concept that is difficult to measure: appropriateness[2]. In common language, an appropriate choice is one which is suitable or proper in the circumstances, but this is surprisingly different from the medical definitions[2].

In an attempt to address this difference when adopting the concept to cardiac imaging, the Appropriateness Criteria Working Group (ACCF AUC Task Force) [28] defined an appropriate test as "*one in which the expected incremental information, combined with clinical judgement, exceeds the expected negative consequences (risks of the procedure i.e. radiation or contrast exposure and the downstream impact of poor test performance, such as delay in diagnostic (false negatives) or inappropriate diagnosis (false positives)) by a sufficiently wide margin for specific indication that the procedure is generally considered acceptable care and a reasonable approach for the indication*" [2, 30]. Because of the low risk of non-invasive imaging, there are many circumstances where this definition could be deemed as insufficient – the risk is almost zero, so the balance of benefit and risk is positive, but the information obtained is still inadequate to justify the performance of the test[2]. A new definition overcomes these concerns by framing the decision in the context of a consensus about "reasonable care"[28], and resource utilisation: "*The concept of appropriateness, as applied to health care, balances the risk and benefit of a treatment, test, or procedure in the context of available resources for an individual patient with specific characteristics*"[2, 28]. Importantly, it is now acknowledged that Appropriate Use Criteria should provide guidance to supplement the clinician's judgment, rather than being prescriptive[2]. While the risk of harm with inappropriate intervention is an important motivator to the application of AUC, the focus on appropriate use in imaging is mainly rooted in resource utilisation and medical expenditure[2].

The first AUC (for Single-photon Emission Computed Tomography SPECT) were launched at the end of 2005, and the first transthoracic (TTE) and transoesophageal (TEE/TOE) echocardiography AUC document was released two years later [2, 31, 32]. Stress echocardiography (SE) AUC was not developed until 2008, therefore not included in the first version of the

echocardiography AUC [33], but the criteria were merged in the updated 2011 version [26]. Previous versions of AUC for the different cardiac imaging techniques were also redesigned and re-evaluated, due to several common clinical scenarios which were not included in the first editions. The AUC continue to evolve, and criteria for multimodality cardiac imaging and the re-definition of "inappropriateness" are counted amongst recent changes [2, 28, 34].

The concept of appropriate use has had an extraordinary influence on the relationships between patients, physicians, administrators and insurance companies over the last decade[2]. However, the most important step commenced January 1, 2017: All health professionals and medical practices must certify their use of appropriate use criteria when ordering advanced diagnostic imaging, as payments under the Medicare & Medicaid Services program in the US[3] will no longer be made for 'inappropriate' tests.

However, new research studies have shown that some tests labelled "inappropriate" have provided significant results leading to a change in management in the patient [35-37].

Disadvantages of the Appropriate Use Criteria

While the AUC have become a foundation of the efforts to improve quality and control expenditure in the US, their uptake in other jurisdictions has been less enthusiastic due to some challenges[2, 35, 38-63]:

1. The AUC differ importantly from clinical guidelines in that they are developed by consensus[30] based on a modified Delphi process which has important disadvantages[1, 64]. The AUC are situation-specific though not necessarily evidence-based. The scientific evidence basis of some AUC is weak as they are based on levels of evidence B or C. A potential question, therefore, is whether they are necessarily correct[1]. This is especially problematic when the only difference between an appropriate or inappropriate test is the time of evaluation or the symptom status of the patient[1] (Table 1- 1).

Table 1- 1 Comparison between some guidelines for the clinical application of echocardiography indications and their corresponding AUC indications (Reproduced from Fonseca, et al.[1])

ACC/AHA Guidelines for the Clinical Application of Echocardiography	Class	AUC for echocardiography	AUC Score
Reevaluation of asymptomatic patients with severe stenosis	I	40. Routine surveillance (< 1 y) of moderate or severe valvular stenosis without cange in clinical status or cardiac exam	I (3)
		41. Routine surveillance (≥1 y) of moderate or severe valvular stenosis without cange in clinical status or cardiac exam	A (8)
Routine reevaluation of asymptomatic patients with mild to moderate mitral stenosis and stable physical signs	III	38. Routine surveillance (< 3 y) of mild valvular stenosis without change in clinical status or cardiac exam	I (3)
		39. Routine surveillance (≥ 3 y) of mild valvular stenosis without change in clinical status or cardiac exam	A (7)
Reevaluation of asymptomatic patients with severe regurgitation	I	45. Routine surveillance (< 1 y) of moderate or severe valvular regurgitation without cange in clinical status or cardiac exam	U (6)
		46. Routine surveillance (≥ 1 y) of moderate or severe valvular regurgitation without cange in clinical status or cardiac exam	A (8)
Reevaluation of patients with mild to moderate regurgitation with ventricular dilation without clinical symptoms	I	43. Routine surveillance (< 3y) of mild valvular regurgitation without cange in clinical status or cardiac exam	I (2)
		44. Routine surveillance (≥ 3 y) of mild valvular regurgitation without cange in clinical status or cardiac exam	U (4)
		45. Routine surveillance (< 1 y) of moderate or severe valvular regurgitation without cange in clinical status or cardiac exam	U (6)
		46. Routine surveillance (≥1 y) of moderate or severe valvular regurgitation without cange in clinical status or cardiac exam	A (8)

In some situations (asymptomatic non-severe valve disease), the test could be expected to have limited or no prognostic value, irrespective of timing and appropriateness[2]. Contrarywise, the mandate of symptomatic criteria to define appropriateness for example in hypertensive heart disease, may allow heart failure events to occur even after tests labelled as inappropriate[2].

2. AUC represent a compilation of indications, but not all situations in which an echocardiogram could be performed are listed in the criteria[1]. Although some studies

of AUC indicate all tests to have been classified [35, 39, 40, 46, 52, 57, 58], in reality, as several indications are often present in the same patient, any non-identified scenario may remain unclassified as secondary reasons for requesting the test are nominated.

3. Several recommendations for echocardiography in current practice guidelines (not just in echocardiography but for disease entities) lack counterparts in the AUC[2]. For example, a class I recommendation is given for follow-up or surveillance after surgery of masses known to have a high likelihood of recurrence (e.g., myxoma [65]). The AUC classification of "suspected cardiac mass" – or even screening – does not cover the described scenario[2].
4. The lack of follow-up events does not itself define inappropriateness[1]. We should expect three outcomes from testing – that the test should provide benefit, that it should reclassify risk, and that the risk should be alterable with intervention. While the link between a low risk for future cardiovascular events and unnecessary testing of patients with a low probability of coronary artery disease is readily understood, this is an imperfect example of inappropriate testing[1]. The original interpretation of "appropriateness" was a situation where the benefit of the test result exceeded the risk (to which we might add financial burden) associated with testing. A gratifyingly negative echocardiogram that reassures an anxious patient and confirms the correctness of a conservative strategy is perhaps the most "appropriate" of scenarios[1]. Conversely, a positive test for ischemia that predicts an adverse event in a patient who is unsuitable for intervention is actually inappropriate. It is the reclassification of risk assessment that provides the most return on the investment of testing[1].
5. The application of AUC to patient selection may be problematic as an audit tool[2]. When an appropriate indication is required to order the test at point-of-service (the facility or location where the test is being performed), the referring clinician may list a co-existing appropriate indication rather than the actual clinical problem (which may be inappropriate)[2]. This is particularly likely when the proportion of inappropriate tests is assessed as part of the echocardiography accreditation process [66]. Additionally, the application of AUC is bedevilled by the reliability and reproducibility of the assessment of appropriateness[2].
6. Retrospective auditing may be especially challenging. For example, a false positive may arise when a test is labelled as "inappropriate", being "appropriate" due to the

possible lack of information that auditors have at the time. The reason for requesting an echocardiogram is often inadequately detailed in the medical records[2].

7. After almost a decade of this work, there are concerns about the actual impact the AUC has had on physician ordering behaviour[1, 2]. Although there has been a steady decrease in the cardiac imaging growth (excluding Cardiac Computed Tomography and Cardiac Magnetic Resonance [67]) since 2006 (more than 15% of the Medicare reimbursement between 2006 and 2011)[25, 67], we see very little evidence of any real impact on the decline of "inappropriate" tests despite continuous attention to the problem of improper use[2]. The literature gives the impression that a similar proportion of "inappropriate" testing, despite experience, educational campaigns and close follow-up continues. Moreover, the correlation between appropriateness and clinical impact has not been well studied [35]. Analysis of the six most relevant studies for assessment of the 2011 AUC for stress echocardiography [44, 50, 68-71], shows that the average of appropriate use is around 61% and that there is no significant correlation between the rate of appropriateness and the enrolment years of the studies[1]. Indeed, the work of Bhattacharyya et al[72] provides a scenario for appropriate selection of stress echocardiography that is neither different nor encouraging[1]. The study shows that 62.4% of the stress echocardiograms were "appropriate", 28.4% were "inappropriate", and 9.2% were "uncertain". Interestingly, these results are similar to the results obtained by Cortigiani et al.,[68] who found that 63.4% of SE were "appropriate", 27.3% inappropriate and 9.3% uncertain under the same AUC. However, with a noticeable difference: the SE evaluated by Cortigiani et al.,[68] were performed between January 2001 and December 2007, and the SE assessed in the study by Bhattacharyya et al., were conducted between October 2010 and September 2011[1]. These results should lead us to focus on the actual impact of Appropriate Use Criteria on health professionals throughout this time. While interventions directed towards interns and residents at the Massachusetts General Hospital,[73] and the development of software to link AUC to ordering in the electronic medical record are promising,[74] it remains uncertain as to whether the processes merely change the attribution of studies based upon these new rules. Certainly, the effects of teaching interventions have not been uniformly favourable, with a prominent study of an intervention based on lectures and training returning negligible results.[1, 70]

8. New studies have shown that some tests labelled "inappropriate" have led to a change in management (around 20%) due to new significant findings (~30%)[35-37, 75]. This is particularly important because the AUC process aims to avoid the performance of labelled "inappropriate" requests in order to control cardiac imaging use and health expenditure. Therefore, by following the AUC, physicians may not discover key findings leading to inappropriate patient care.

Although there is an increased interest in adopting the AUC to the Australian practice, the application of AUC remains primarily a North American phenomenon. There are many aspects of medical care in the USA which are mysterious to foreigners. Are AUC merely a component of this difference, or is it time to declare victory for AUC and apply them more widely than in the USA?[1]

The listed disadvantages of the AUC should caution us regarding more widespread adoption of AUC[1]. There is no doubt that we need a method to reduce the over-use of cardiovascular imaging. The AUC appear to be a part of the imaging landscape in North America, but it may not be the solution to the challenges of imaging selection elsewhere[1]. Perhaps those of us outside the USA should contemplate a metric that better measures the information content of testing, in order to find the desired balance between clinical utility and reduction of health expenditure. Judging appropriateness in echocardiography is a process based on knowledge, experience, information, resources and the real desire to provide an adequate service to the patient. It does not sit well with formulaic approaches based on uncritical applications of AUC[1]. Importantly, physicians should be formulating their own conclusions regarding test requirements and using the AUC as a supplement to this rather than the sole arbitrator[28].

Hypothesis and aims of this thesis

The disadvantages and limitations described previously are unfortunate as AUC may have significant health economic implications. Due to those issues, the overall objectives of this thesis were first to examine the use of cardiac imaging in Australia and secondly, if the AUC are suitable to use in this country in the event Australia needs an appropriate use improvement program. However, taking into account the lack of transparency and risks that the AUC process entails when used at the point-of-care/request (the location where the test is requested) due to the possible change of indications and the risk that cancelling the labelled "inappropriate" requests has for patients, we also sought to find an alternative process which can be used at the

point-of-service in an easy and efficient way to avoid inappropriate testing with the safety of the cardiac evaluation. In order to accomplish these objectives, the following are the hypothesis and specific objectives of this thesis.

Hypothesis: the AUC are an inconclusive tool to improve appropriateness and use of cardiac imaging.

Objectives:

1. To define the use of cardiac imaging in Australia and to analyse related factors in order to establish if a program for appropriate use is needed in this country.
2. To identify if the actual AUC are suitable for use in Australia by addressing the following objectives:

    a. To determine differences between AUC and guidelines.
    b. To determine the trends of the percentage of appropriateness in the different cardiac imaging.
    c. To determine the impact of the AUC in doctors' ordering behaviour.
    d. To determine if the concept of appropriateness used by the AUC affects health outcomes.
    e. To define factors that influence ordering behaviour in clinicians.
3. To propose an improvement plan to avoid possible inappropriate requests without the risk of missing significant cardiac findings.

Structure of the thesis

Chapter 1. Introduction

Chapter 2. Methodology

Chapter 3. Appropriate use of imaging in Australia.

Chapter 4. Evidence status of the Appropriate Use Criteria.

Chapter 5. Impact of Appropriate Use Criteria on clinician's behaviour.

Chapter 6. Impact of Appropriate Use Criteria on health outcomes: Heart Failure Survival

Chapter 7. Understanding Cardiac Imaging Decision-Making: Appropriate Use Determinants

Chapter 8. Identification of inappropriate request at the point-of-service

Chapter 9. Usefulness of hand-held echocardiography in inappropriate requests

Chapter 10. Summary and discussion

Concluding remarks

This chapter provided an overview of the AUC, their motivation and aims. An overview of the findings suggests that the AUC have important limitations that should be addressed before adoption for the use of the criteria outside the USA.

Knowing the characteristics and use of cardiac imaging in Australia and understanding the challenges of the AUC provides us with the opportunity to determine if the criteria are suitable for use in this environment.

This thesis aims to evaluate the potential problems of using the American AUC and to identify a solution that avoids inappropriate tests in Australia. Findings of this research will fill the current evidence gaps.

The following chapter describes the methodology used for this research.

Chapter 2. Methodology

Preface

In the previous section, background information was provided regarding the significant growth of national health care budgets globally and the recent shift to understand how the imaging component can be spent wisely.

It was identified that cardiovascular imaging has seen the most significant growth rate in the last two decades in the USA. The need to utilise a method for determining appropriateness of imaging use and controlling the use of cardiac imaging services was discussed. As a response to this need, it was shown that the Appropriate Use Criteria for different cardiac imaging techniques were developed. However, it was considered that these criteria have some disadvantages and flaws, which interfere with the ability of the criteria to impact clinician's behaviour to achieve appropriate use of imaging. Moreover, the intention to use the criteria broadly outside the USA, including Australia, has been the basis for this research.

The present chapter briefly describes the methodology employed in this thesis. This thesis comprises quantitative and qualitative research methods. Additional factors that are unique to each chapter are described in more detail within the methodology section of each of the subsequent chapters. More details related to particular investigations are addressed in their respective studies that are presented in chapters 3 to 9.

This research was framed by the Tasmanian Health and Medical Human Research Ethics Committee approvals H0014017 and H0015516.

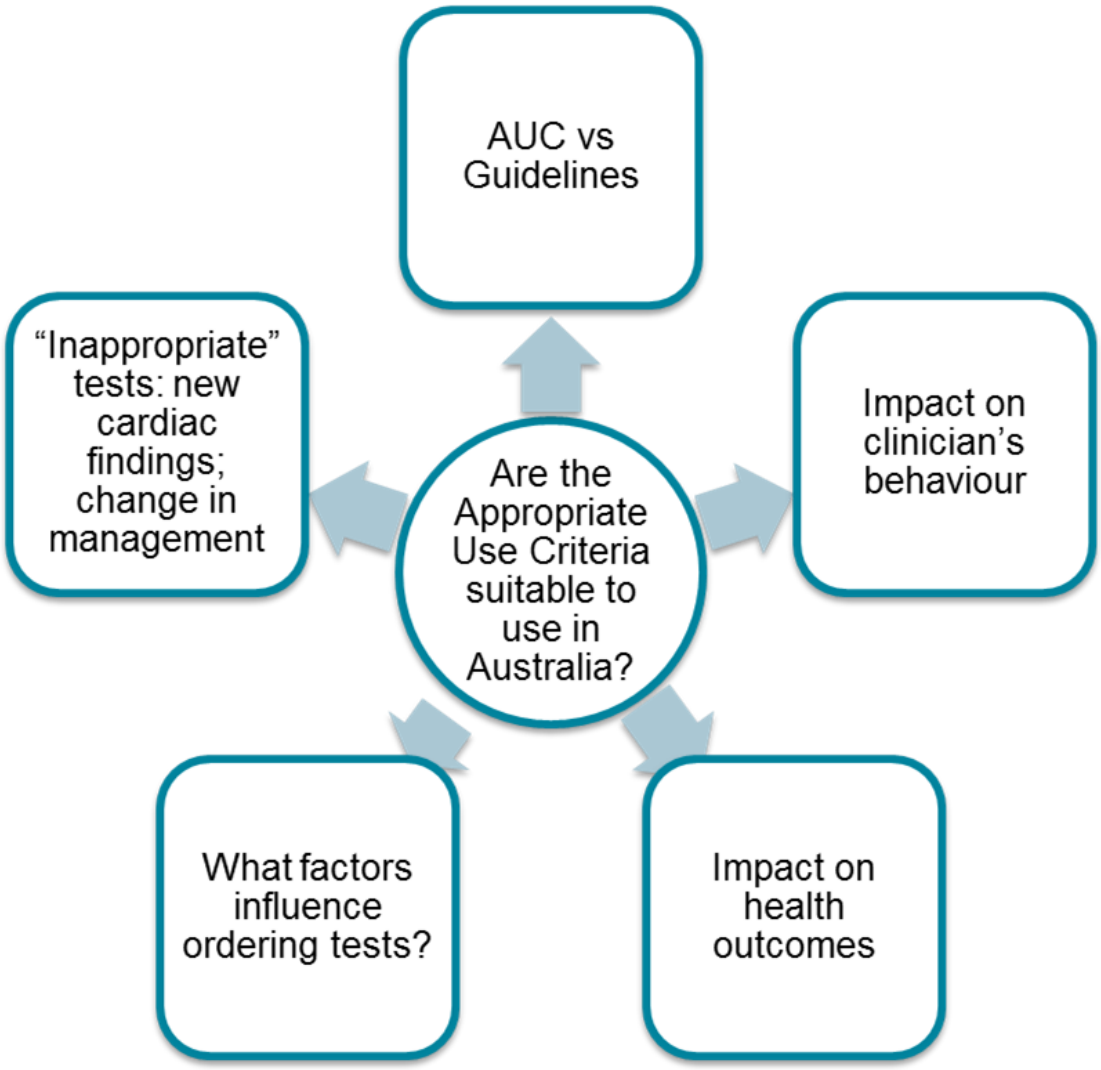
Methodology

This thesis was developed in three stages. The first phase, contained in Chapter 3, determined how cardiac imaging is used in Australia to define the need (or not) of appropriate use. Two areas were analysed. First, the growth of cardiac imaging use in Australia and second, the associations of imaging use with demographic characteristics, availability of medical workforce and cardiovascular burden of disease for each Medicare Locals[4]. Data from Medicare Australia Statistics, Australian Health Survey, and Health Workforce were analysed to define the growth and regional variation in the use of cardiac imaging in the last decade. The Medicare Statistics website contains quarterly and annualised data (financial and calendar years) relating to Medicare [76].The increase of imaging was determined using use of tests per 100,000 people in each geographical area over time intervals. Medicare locals were chosen as the regional areas for analysis due to them being the smallest geographical areas with available information. Spearman correlations, negative binomial regression and zero-inflated negative binomial regressions, were used to investigate univariable and multivariable associations between age-weighted testing and characteristics of each Medicare local[4].

The second stage addressed the disadvantages and challenges that the criteria have faced during the last decade to determine if the AUC are the correct tools to use in Australia. This phase was developed through Chapters 4 to 7.

The core research question was: Are the AUC suitable to use in this country? Figure 2- 1 shows the topics that were taken into account to develop this aim.

Figure 2- 1 Topics



The matters to investigate were the disadvantages and challenges of the criteria that were explained in the Introduction. In order to understand the evidence status of Appropriate Use Criteria (Chapter 4), a comparison between the criteria and the published Cardiology guidelines was performed to find concordance between both[5]. Concordance was determined when "appropriate" items in the criteria had recommendations class I or IIa in the guidelines, "inappropriate" scenarios had recommendations class III, and "uncertain" items had recommendations class IIb[5]. Figure 2- 2 shows the definition of recommendations in clinical guidelines[77].

Figure 2- 2 Guidelines' recommendations (Adapted from Gibbons, et al.[77])

As shown in Figure 1-4, the appropriateness designation and guidelines recommendations must be consistent.

The second topic addressed to determine the suitability of AUC use was an investigation regarding the temporal changes in the appropriateness of cardiac imaging, documenting the impact of the criteria on clinicians' ordering behaviour. This investigation determined the change of proportion of "appropriate/uncertain/inappropriate" tests over time, which implies impact on requesting behaviour. This exploration is included in Chapter 5 which was a systematic review of published manuscripts related to the AUC use in cardiovascular imaging. Two reviewers conducted the literature search of online databases. Papers were included in the analysis if they specified type and edition of AUC used, year of data collection, sample size of tests evaluated, proportion of "appropriate" and inappropriate or uncertain studies and proportion of "classified" studies, which correspond to the percentage of studies whose reasons were found in the list of criterion of the AUC[6]. The proportions of "appropriate", "uncertain", and "inappropriate" testing over time were analysed in meta-regressions using logit transformation to calculate the weighted summary proportion with random-effect models. Ten pooled analysis were conducted for each cardiac imaging found and edition of AUC used: TTE using 2007 AUC, TTE using 2011 AUC, TEE using 2007 AUC, TEE using 2011 AUC, SE using 2008 AUC, SE using 2011 AUC, CCT using 2006 AUC, CCT using 2010 AUC, and SPECT using 2005 and 2009 AUC. Publication bias was examined using funnel plots and Egger's test[6].

The impact of the criteria on health outcomes in Chapter 6 was an analysis of event-free survival, cause-specific hazard and cumulative incidence of Heart Failure (HF) readmission and death of patients with HF utilising Kaplan-Meier and Cox proportional hazard regression and competing risk analyses[7]. This was a cohort study of patients hospitalised for HF at a tertiary referral hospital. Patients were included if they were adults (≥18 years old), had a

diagnosis of HF and a TTE before or during hospitalisation. Participants were divided into four groups according to the definitions of follow-up TTE in the AUC for echocardiography[26, 78]: patients with no follow-up TTE after discharge, patients with "inappropriate" follow-up TTE according to AUC, after discharge (follow-up <1 year with no change in symptoms or clinical status), patients with "uncertain" follow-up TTE according to AUC after discharge (>1 year with no change in symptoms of clinical status) and patients with "appropriate" follow-up TTE according to AUC (TTE due to change in symptoms or clinical status)[7].

Chapter 7 is a qualitative analysis performed using semi-structured face-to-face interviews with open-ended questions to analyse factors leading doctors to order echocardiograms and the relationship with inappropriate testing. Cardiologists and non-cardiologists were included in the study.

The third stage of this work was to provide a solution to avoid the performance of tests labelled "inappropriate" in a safe way. This phase was based on the findings of the previous studies, which demonstrated inconsistencies between clinical guidelines and AUC, nominal impact on the request of inappropriate testing over time, a lack of clear effect on event-free survival, and a mismatch between the clinical reasoning process of doctors and the AUC. These results with the assumption that doctors can choose "appropriate" indications instead of the presenting problem of the patient (which could be classified as "inappropriate"), and that some tests labelled "inappropriate" have provided information leading to a change in patient management, gave cause to a further two studies which were developed to avoid these issues. These studies are included in chapters 8 and 9.

In chapter 8, the objective was to find a transparent way to determine labelled "inappropriate" tests at the point of service (e.g. echo-laboratory). This approach aimed to nullify the possibility of shifting indications at the point of request by practitioners. General characteristics of inappropriate TTE were determined and summarised in a questionnaire of four binary questions most commonly associated with "inappropriate" tests[8]. These questions are related to the absence or change of cardiovascular symptoms or signs, routine surveillance as the purpose of the test, previous tests in the last year and if endocarditis with no murmur or positive blood cultures was the reason for the scan[8]. These questions were applied to two different validation groups. Analyses of specificity, sensitivity, and predictive values were performed to determine the accuracy of the prediction of inappropriate requests.

In Chapter 9, two approaches to "inappropriate" requests (standard TTE versus hand-held echocardiography) were compared in a case-control study. Patients were followed for a period of 6 months and outcomes such as time until the scan, repeat echocardiography, the cost of each strategy, new findings and impact on management were determined[9].

Statistical analyses in all chapters were performed using R software[79]; geo-mapping for Chapter 2 was done using QGIS 2.4.0 (QGIS 2.4.0, Open Source Geographic Information System; OpenSource Geospatial Foundation, Beaverton, OR, USA); hand-held ultrasonography was performed using the GE Vscan V 1.2 hand-held device and standard TTE was performed GE Vivid 9, Philips IE33 and Acuson SC2000[9].

The following chapter addresses the first objective of this thesis: the growth in use of cardiac imaging in Australia and related factors of use in order to understand the 'appropriateness use' status in this country.

Chapter 3. Appropriate use of cardiac imaging in Australia

The research contained within this chapter has been published as [4]:

- Fonseca R, Otahal P, Wiggins N and Marwick TH. Growth and geographical variation in the use of cardiac imaging in Australia. *Intern Med J*. 2015;45:1115-27.

Preface

In Australia, health expenditure has increased from $50.3 billion (6.5% of GDP) in 1990 to $154.6 billion (9.7% of GDP) in 2014[80, 81]. Moreover, cardiovascular disease (CVD) has the highest level of health-care expenditure of any disease group[80, 81].

In 2015, the Medicare reimbursement for TTE only was approximately $186.0 million Australian dollars[76]. Assuming that low-value tests oscillate between 7% and 20% (according to international literature), the Medicare reimbursement in Australia would be expected to fluctuate between $13.0 and $37.2 million dollars[8, 76].

In this country, several studies have shown a wide difference in the use of healthcare services by geographical areas indicating ineffective use [80-82]. The key factors of increases in use and local variations in the usage of health services are possible measurements of quality of care and appropriate use[12].

In the previous chapter, general information was provided on why the Appropriate Use Criteria (AUC) were needed to control the use of cardiac imaging in the USA. However, despite the increased interest in using the same criteria within the Australian practice, the potential impact and relevance of the AUC have not been assessed in this country.

The present chapter aims to determine the growth in cardiac imaging use in Australia during the last decade and most importantly, to establish the associations of imaging use with demographics and burden of disease in order to analyse potential misuse of this health resource.

Abstract

Growth rates and regional differences in the use of cardiac imaging are potential metrics of quality of care. This study sought to define growth and regional variation in outpatient cardiac imaging in Australia.

Methods.

Analyses are based on the rate of outpatient transthoracic (TTE), trans-oesophageal (TOE) and stress echocardiography (SE) and single-photon emission computed tomography (SPECT) per 100,000 people in each geographic insurance region in Australia (Medicare Local, ML). Numbers of tests from 2002-13 were obtained from Medicare Australia Statistics, and the number of doctors was obtained from the Health Workforce data. Demographic information (total population, rural areas, and quintiles of disadvantage) were obtained from census data.

Results.

Over the last eleven years, TTE reimbursements/100,000 people increased from 1,780 to 3,497 (8.8% annualised growth), TOE from 33 to 61, SE from 181 to 947 and SPECT from 287 to 337. SE had the biggest incremental growth, with an average rate of increase at 38.5% per year. The relationships between the use of each cardiac imaging technique and demographic, medical and illness factors were explored through outpatient tests reimbursed in 2012. It was found that for each additional medical practitioner per 1,000 people, there was an increase of 1.25 times the rate of TTE (95% CI: [1.17, 1.33], $p<0.001$, and TOE use (B=1.13 [1.04, 1.24], p=0.005), independent of regional burden of cardiovascular disease and social determinants. For SPECT the largest independent correlate for testing was the percentage of women within the ML; each additional percentage increase resulted in doubling of the rate of testing (B=2.25 [1.72, 2.94], $p<0.001$).

Conclusions.

Variation in the use of TTE in Australia is not illness-related and may be evidence of under and over-utilization. An appropriate use process may contain this variation.

Keywords

Clinical Practice Variation, Echocardiography, SPECT, Misuse of Health Services, Healthcare Costs

Introduction

Increased life expectancy, population growth - especially of the elderly - and access to new treatments and technology have all contributed to the rise in costs and demand in the Australian health system over the last decade [80, 83].

Cardiovascular disease continues to be associated with the highest level of health-care expenditure [81], and has been a major contributor to this growth. Prior to a recent reduction in the pace of imaging growth in the USA [25, 67, 84], imaging costs had exceeded the growth of many other medical activities [18]. Cardiovascular imaging was a significant contributor to that process. Besides the increase in demand of imaging services, studies have shown a wide difference in the use of healthcare services by geographical areas that indicates ineffective use[85, 86].

Geographical differences in resource utilisation provide a potential means to understanding the causes of variations in care [82, 87, 88]. Such an analysis can identify the relationship between demographics, disease burden, and access to care [24]. Available data about regional variation in the use of echocardiography [24, 89] includes USA data from the Dartmouth Atlas (which preceded the development of appropriate use criteria (AUC) and the reduction in the rate of imaging growth) and local echocardiography numbers in the UK [89].

As there are a number of differences in care delivery in Australia (including universal health cover, remoteness and access)[87] we sought to analyse the growth of the use of cardiac imaging as well as the causes of variation of use in this country.

Methods

Study design.

We used available Medicare statistics between2002 and 2013 to measure trends in the use of outpatient transthoracic (TTE), transesophageal (TOE) and stress echocardiography (SE) and single-photon emission computed tomography (SPECT) per 100,000 Medicare beneficiaries in Australia. The 2011-2012 Australian Health Survey provided data, which permitted a cross-sectional study of regional associations of cardiac imaging performed in the 2012 calendar year. The geographic units of measurement were Medicare Locals (ML), made up of 61 regional primary health care organisations covering the whole of Australia, which were developed to connect local health services better [90]. Health utilisation statistics are available for each ML and statistics from other sources can be mapped to these regions. As MLs differ in population, health status, demographics, remoteness and socioeconomic levels, this study additionally

used seven clusters of MLs called peer groups (Metro 1, Metro 2, Metro 3, Regional 1, Regional 2, Rural 1 and Rural 2) taking into account proximity of each ML to a metropolitan area, proximity to large hospitals and socioeconomic characteristics as described by the National Health Performance Authority [87].

Data sources.

Data were obtained from Medicare Australia Statistics, the Australian Bureau of Statistics (ABS) and ABS Health Survey and Health Workforce Data.

The numbers of imaging tests were extracted from the Medicare Australia Statistics,[76] using the cardiac ultrasound codes 55113, 55114, 55115, 55119, 551120, 55121 for TTE; 55118, 55125 for TOE; 55116, 55117, 55122 , 55123 for SE and 61307, 61654 for SPECT [91]. In order to compensate for the age-dependent nature of many illnesses requiring cardiac imaging, the rate of use of each modality per 100,000 people was age-weighted to reflect the age of each region relative to 37 years (the mean age for Australia)[92].

Demographic data (total population, rural areas, and quintiles of disadvantage) were obtained from the "Estimated Resident Population" statistics for MLs for 2012 [93]. Areas were categorised as most socioeconomically disadvantaged if they were in quintiles 1, 2, and 3 of the Socioeconomic Index for Areas (SEIFA) scores. We defined "metropolitan area" as the MLs listed in the Metro 1, 2 and 3 groups, regional area as MLs in the Regional 1 and 2 peer groups and rural area as the MLs in Rural 1 and 2 peer groups. For multivariable models, metro areas were compared with grouped regional and rural areas.

Medical workforce information was extracted from the National Health Workforce Dataset, a national database [94]. The number of "non-cardiologists" was outlined as the difference between the total number of practitioners and the number of cardiologists in each ML. Data were analysed taking into account the number of non-cardiologists per 1,000 persons instead of the number of cardiac specialists, as the latter group, in most cases, receives a referral and actually perform the tests.

The burden of disease was obtained from the percentage of people with cardiovascular disease (extracted from the ABS Health Survey for persons aged 18 and over[95]), and cardiovascular mortality per 1,000 (from the national death index).

Statistical analysis.

We used Medicare data from 2002 to 2013 to calculate the trend in utilisation of cardiac imaging per 100,000 people. Detailed analyses were performed only on data from 2012 for which most information was available.

Spearman rank correlations were used to investigate univariate associations between age-weighted testing and characteristics of each ML. For TTE and SPECT analyses, we used negative binomial regression to build models for each rate (age-weighted tests per 100,000); SE and TOE outcomes contained many MLs with zero testing, therefore we switched to zero-inflated negative binomial (ZINB) for these analyses using the number of cardiologists to predict the inflation model.

The coefficients for independent variables coded as percentages (e.g. women) are interpreted as a change in rate ratio of age-weighted tests per 1% rise in the variable. Maps were made using standard software (QGIS 2.4.0, Open Source Geographic Information System) and all statistical analyses were made using R 3.2.1.[79]

Results.

Population characteristics.

Appendix Table 1 displays the population, number of doctors and burden of cardiovascular disease in Australian Medicare Locals. Most of the Australian population lives in metropolitan areas, and MLs in or near large urban centres (Metro 1 and 2) have the least disadvantaged people.

In general, metro areas 1 and 2 have the most doctors (cardiologists and non-cardiologists) per 100,000 people. Access to medical practitioners decreases in progressively more remote MLs. Regional Australia (ML peer groups Regional 1 and 2) have the highest mean age and percentage of people >65 years. However, remote areas (Rural area 2) have the lowest mean age and percentage of individuals more than 65 years, as well as the highest proportion of men.

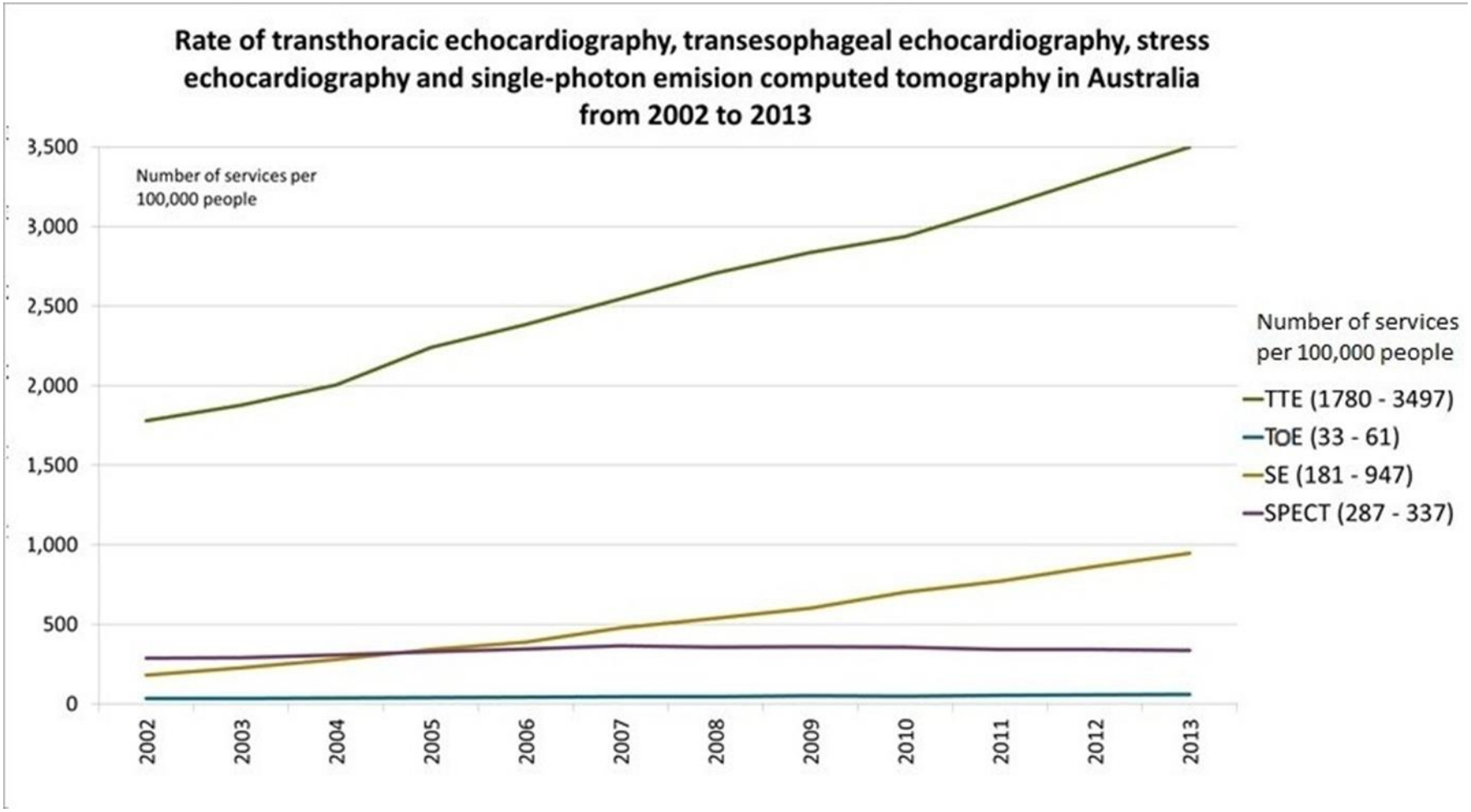
In general, MLs listed as regional or rural areas, have the highest rates of cardiovascular deaths per 100,000 people. The same trend is apparent when cardiovascular disease burden is analysed as the percentage of people >18 years old with CV disease.

Growth of cardiac imaging in Australia.

Medicare Australia reimbursed 973,284 outpatient echocardiograms in the year 2012 (TTE, TOE and SE). Among the cardiac imaging modalities that we are discussing, SE had the biggest change over the last eleven years: a total growth of 423% with an average yearly growth rate of 38.5%.

TTE had an increase of 96% during this period (an average growth per year of 8.8%). SPECT showed an initial rise followed by a relative decrease of its growth rate after 2006. However, there was a rise of 17% in the rate of SPECT per 100,000 people compared to eleven years ago. The upsurge in testing volume over the last eleven years is detailed in Figure 3- 1.

Figure 3- 1 Increment in testing volume in Australia between 2002 and 2013

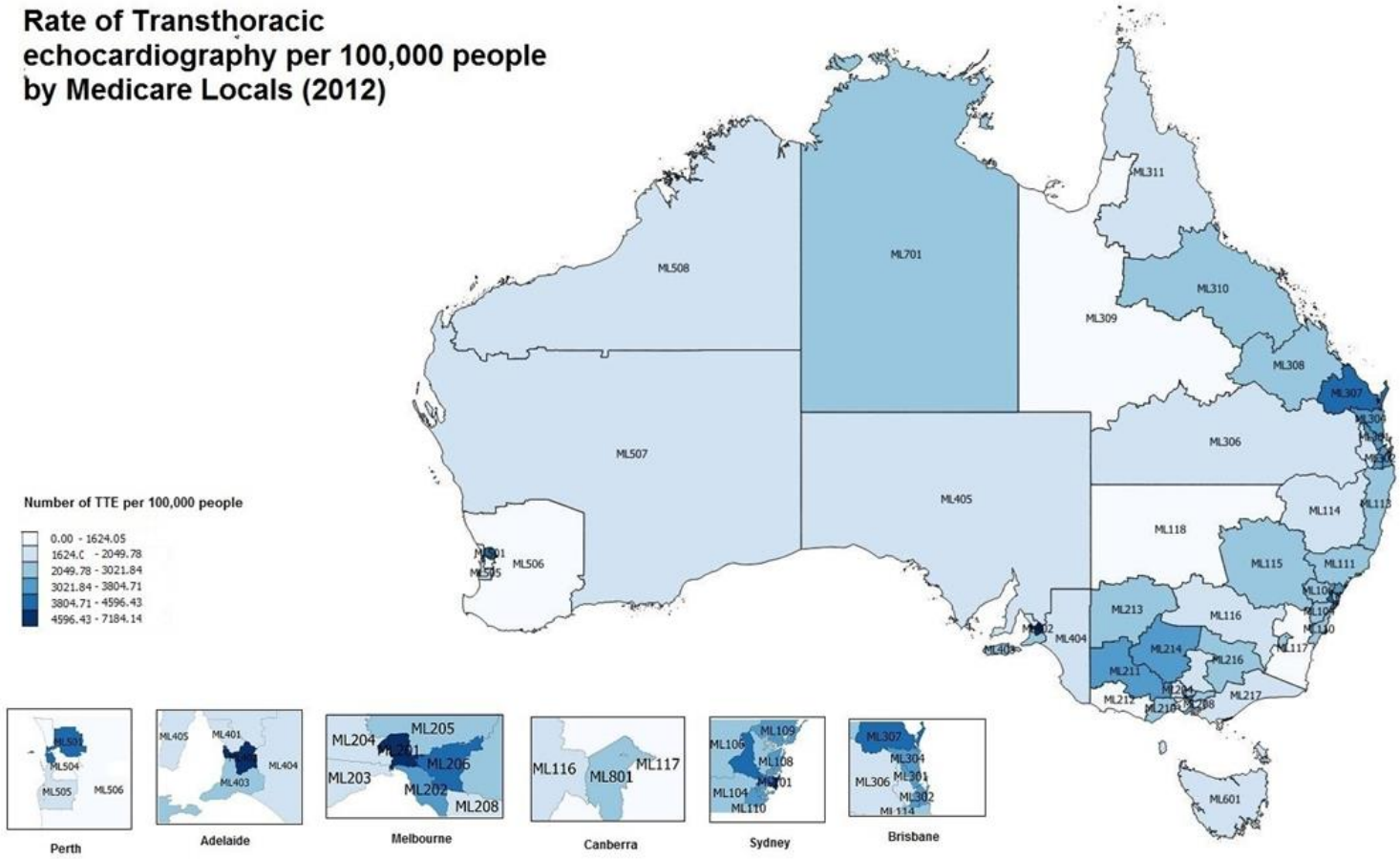


Associations of the regional variation of transthoracic echocardiography.

The national rate for TTE (age weighted) was 3,106 per 100,000 people but the rate varied from 382.8 to 7,184 per 100,000 people, an 18.8-fold variation among regions (Appendix Table 2).

The highest rates of use were found in central metropolitan regions (Inner North West Melbourne, Eastern Sydney and Central Adelaide and Hills, ranging between 6,831 and 7,184 per 100,000 people). The lowest rates were mainly found, but not exclusively, in rural areas (Central and North West Queensland, South West Western Australia, Perth North Metro, and Bentley – Armadale; range 383 – 806 tests per 100,000 people). Far West New South Wales was not included in the study as no data were available. (Appendix Table 2, Figure 3- 2).

Figure 3- 2 Rate of transthoracic echocardiography (age-standardized) per 100,000 people by Medicare Locals in Australia in 2012



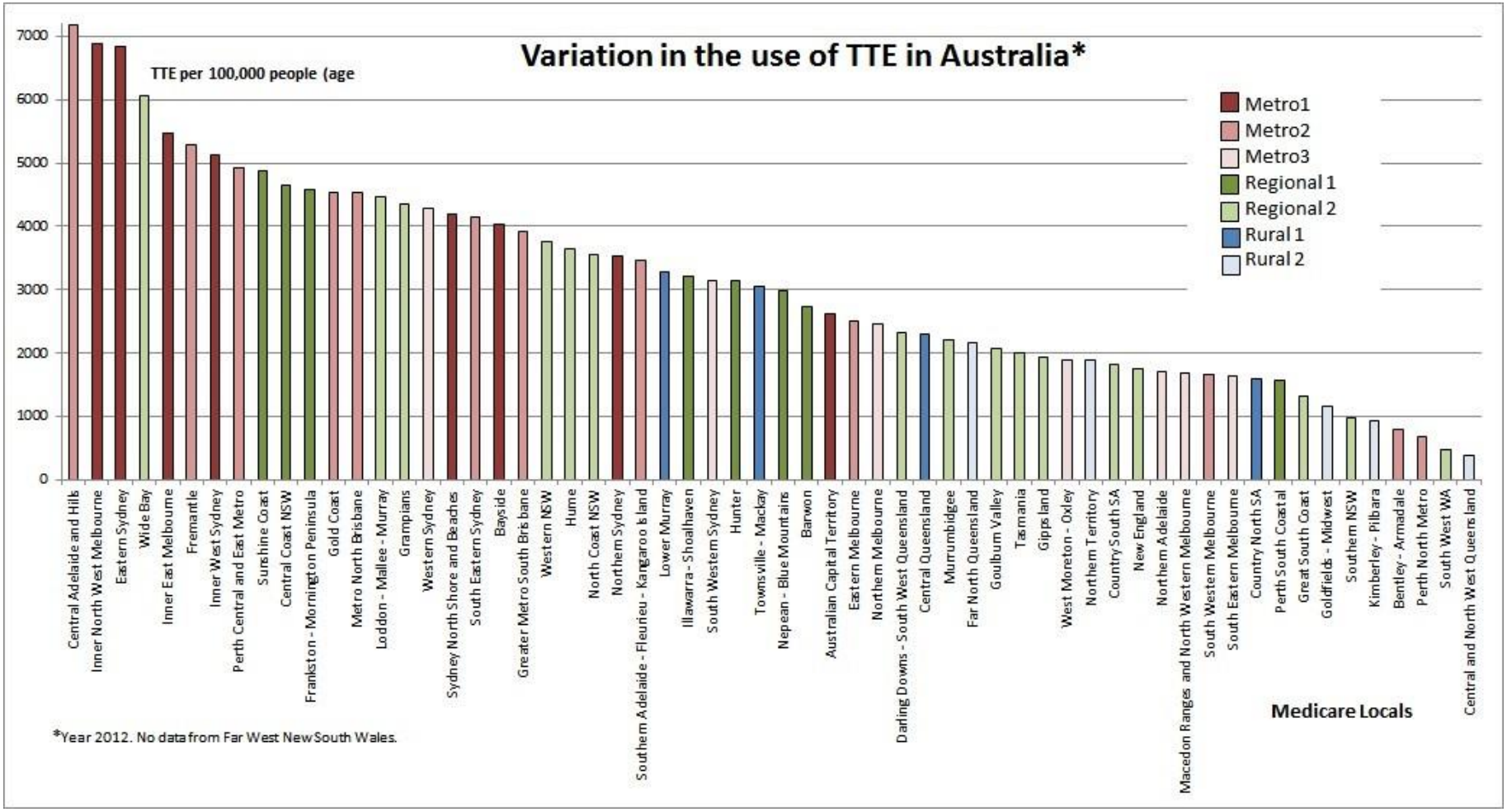
In general, the age-weighted rate of TTE was substantially lower in regional and rural than metropolitan areas, and there was a negative correlation between remoteness and use of TTE (rho= -0.35) (Table 3- 1).

An interesting exception is Wide Bay ML, which has the highest rate of TTE per 100,000 people (6,073) in the Regional 1 peer-group. This rate is higher than some Medicare Locals in metropolitan areas such as Inner East Melbourne, Fremantle or Inner West Sydney (Appendix Table 2, Figure 3- 2 and Figure 3- 3).

Table 3- 1 Correlations between each modality of cardiac imaging and other independent variables

Correlations	**TTE**		**TOE**		**SE**		**SPECT**	
	rho*	p	rho*	p	rho*	p	rho*	p
More disadvantaged (%)	-0.30	0.02	-0.46	0.01	-0.50	<0.01	0.12	0.45
Women (%)	0.59	<0.01	0.33	0.02	0.50	<0.01	0.35	0.03
Non cardiologist (/1,000)	0.75	<0.01	0.65	<0.01	0.63	<0.01	0.14	0.38
Older than 65years (%)	0.32	0.01	0.03	0.86	0.03	0.83	0.49	<0.01
Cardiovascular disease (%)	-0.05	0.71	-0.18	0.22	-0.20	0.13	0.18	0.28
Cardiovascular deaths(/1,000)	0.11	0.41	-0.03	0.83	-0.05	0.72	0.39	0.01
Region(1=Metro/2=Regional/3=Rural)	-0.35	0.01	-0.38	0.01	-0.58	<0.01	0.15	0.36
*Spearman correlation coefficient								

Figure 3- 3 Variation in the TTE use in Australia from the lowest to the highest rate in 2012



Performance of TTE per 100,000 people (age weighted) was positively correlated with percentage of women, Metropolitan MLs, older people and number of non-cardiologist physicians per 1,000. There was no significant correlation between TTE per 100,000 people and percentage of people over 18 years old with cardiovascular diseases or deaths due to cardiovascular causes (Table 3- 1). Additionally, testing was negatively correlated with more disadvantaged people.

The association of these variables in the use of TTE is shown in Appendix Table 3. For each additional 1% rise in the percentage of women within a ML the rates of TTE testing rose 1.22 times [95% CI 1.12-1.34], p <0.001. Furthermore, as non-cardiologists increase by one per 1,000, there is a corresponding increase in rates of testing (β =1.22 [95% CI 1.14-1.31], p <0.001).

Conversely, as the percentage of more disadvantaged people increases the rates tend to decrease (β =0.99 [95% CI 0.99-1.00], p =0.039). In addition, rates of TTE are lower in regional and rural zones compared to metropolitan areas (β =0.80 [95% CI: 0.59, 1.07], p =0.133; β=0.51 [95% CI: 0.34, 0.76], p <0.001, respectively). However, the number of non-cardiologists per 1,000 people and the percentage of women were independently associated with testing, regardless of prevalence of cardiovascular disease, socioeconomic status and metropolitan zone (Table 3- 2).

Interestingly, burden of disease, shown as percentage of people older than 18 years with cardiovascular disease (β=0.99 [95% CI 0.96, 1.02] p=0.420), and cardiovascular deaths (β=1.05 [95% CI 0.87, 1.27] p =0.616), were not associated with testing.

Table 3- 2 Independent associations of population of characteristics in Medicare Locals with age-weighted numbers of tests/100,000 persons

	TTE				**TOE**				**SE**				**SPECT**			
	B	95%CI		p	B	95%CI		p	B	95%CI		p	B	95%CI		p
Women (%)	1.164	1.081	1.254	<0.001	1.096	0.835	1.439	0.509	1.148	0.703	1.875	0.580	2.247	1.720	2.936	<0.001
Metro Medicare Locals	0.960	0.734	1.257	0.768	1.021	0.618	1.686	0.935	1.152	0.547	2.427	0.709	0.409	0.197	0.851	0.017
Cardiovascular disease (%)	1.017	0.995	1.040	0.134	0.979	0.939	1.021	0.332	0.981	0.920	1.046	0.563	_	_	_	_
Cardiovascular deaths (/1,000)	_	_	_	_	_	_	_	_	_	_	_	_	0.896	0.546	1.468	0.662
Non cardiologist (/1,000)	1.247	1.171	1.327	<0.001	1.134	1.038	1.240	0.005	1.127	0.966	1.315	0.129	1.117	0.988	1.262	0.077
More disadvantaged (%)	1.002	0.995	1.009	0.511	0.995	0.984	1.007	0.400	1.000	0.982	1.019	0.973	1.006	0.991	1.020	0.448
Cardiologists (/100,000)*					0.338	0.179	0.637	0.001	0.411	0.255	0.661	<0.001				

* Estimates from inflation part of zero-inflated negative binomial (ZINB) model, only for TOE and SE models.

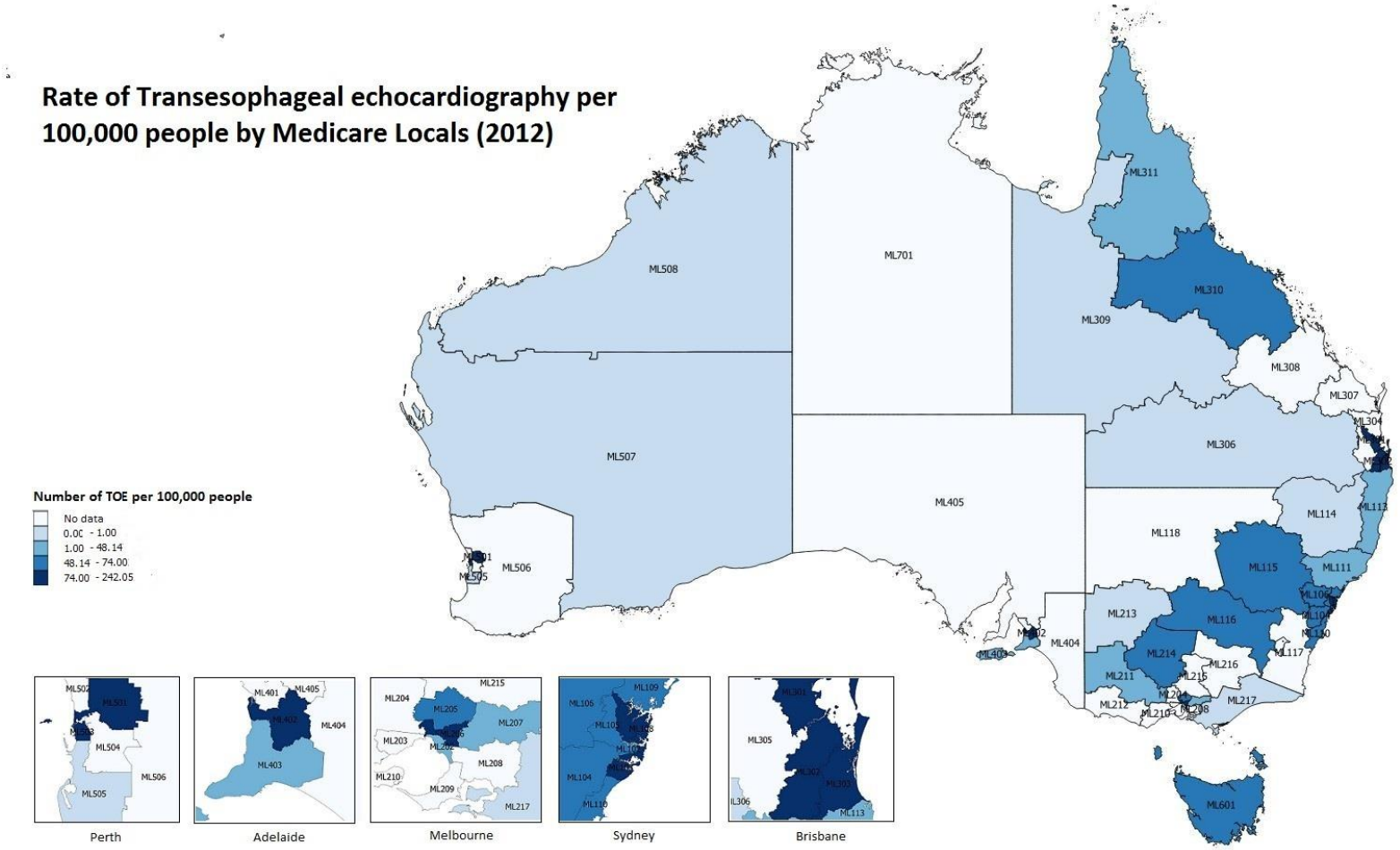
Correlates of the regional variation of TOE.

Data for TOE were available in 50 of 61 Medicare locals (64%), which showed a variation of 0 to 242 per 100,000 (Figure 3- 4). When the lowest rate was excluded, the range of TOE per 100,000 people was 5 to 242, a 46.5-fold variation. The use of TOE was positively correlated with the proportion of women, number of non-cardiologists per 1,000 people, residing in big urban areas and having higher socioeconomic status (Table 3- 1).

The use of TOE was found to be less frequent in regions with a large cardiovascular disease burden (β=0.93 [95% CI 0.91-0.96] p <0.001) or mortality (β=0.77 [95% CI 0.58-1.02] p=0.064), and rates of TOE tests were independently associated with numbers of non-cardiologists (Table 3- 2).

The inflation part of the ZINB regression model shows a decrease of 66% in the odds of a true zero (no testing) as the number of cardiologists increases by 1 per 1,000 people (β=0.34 [95% CI 0.18, 0.64], p =0.001).

Figure 3- 4 Rate of trans-oesophageal echocardiography per 100,000 people (age-standardized) by Medicare Locals in Australia in 2012



Correlates of the regional variation of stress echocardiography.

The rate varied from 0 to 2,992 tests per 100,000 people (Figure 3- 5). Excluding the lowest rate, there was still a 9.6-fold variation for SE (range 311-2,992) (Appendix Table 2).

The use of SE was similarly associated with regions having a larger proportion of women, number of doctors, metropolitan areas and higher socioeconomic status (Table 3- 1), but none of these were shown to be independent associations (Table 3- 2).

Further, there was a decrease in the odds of a true zero (no tests) associated with an increase in cardiologists in each ML (ZINB regression model: B= 0.41 [95% CI 0.26, 0.66] p <0.001).

Correlates of the regional variation of SPECT.

The rate of SPECT per 100,000 people varied from 0 to 955/100,000 in 41 of 61 MLs where data were available for analysis (Figure 6).

There was a 7.2 fold-variation after exclusion of the lowest rates of use (Appendix Table 2). The use of SPECT was positively correlated with percentage of women, percentage of people older than 65 years, and cardiovascular deaths (Table 3- 1).

Cardiovascular disease was not correlated with testing; thus, multivariate analysis was performed using cardiovascular deaths as the indicator of burden of disease. Multivariable analysis showed that proportion of women and metro region were the only independent correlates of imaging use (Table 3- 2).

Figure 3- 5 Rate of Stress echocardiography per 100,000 people (age-standardized) by Medicare Locals in Australia in 2012

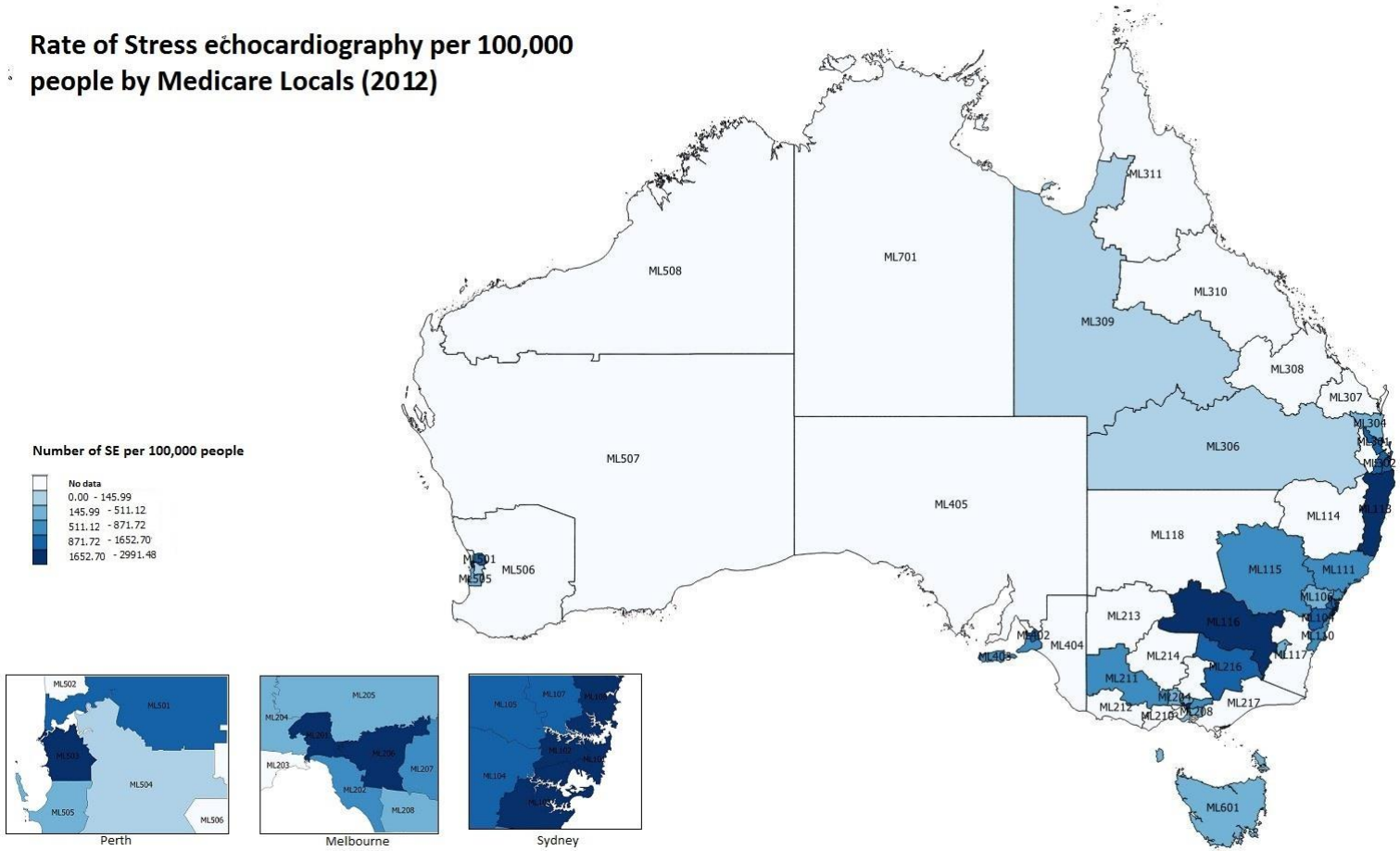
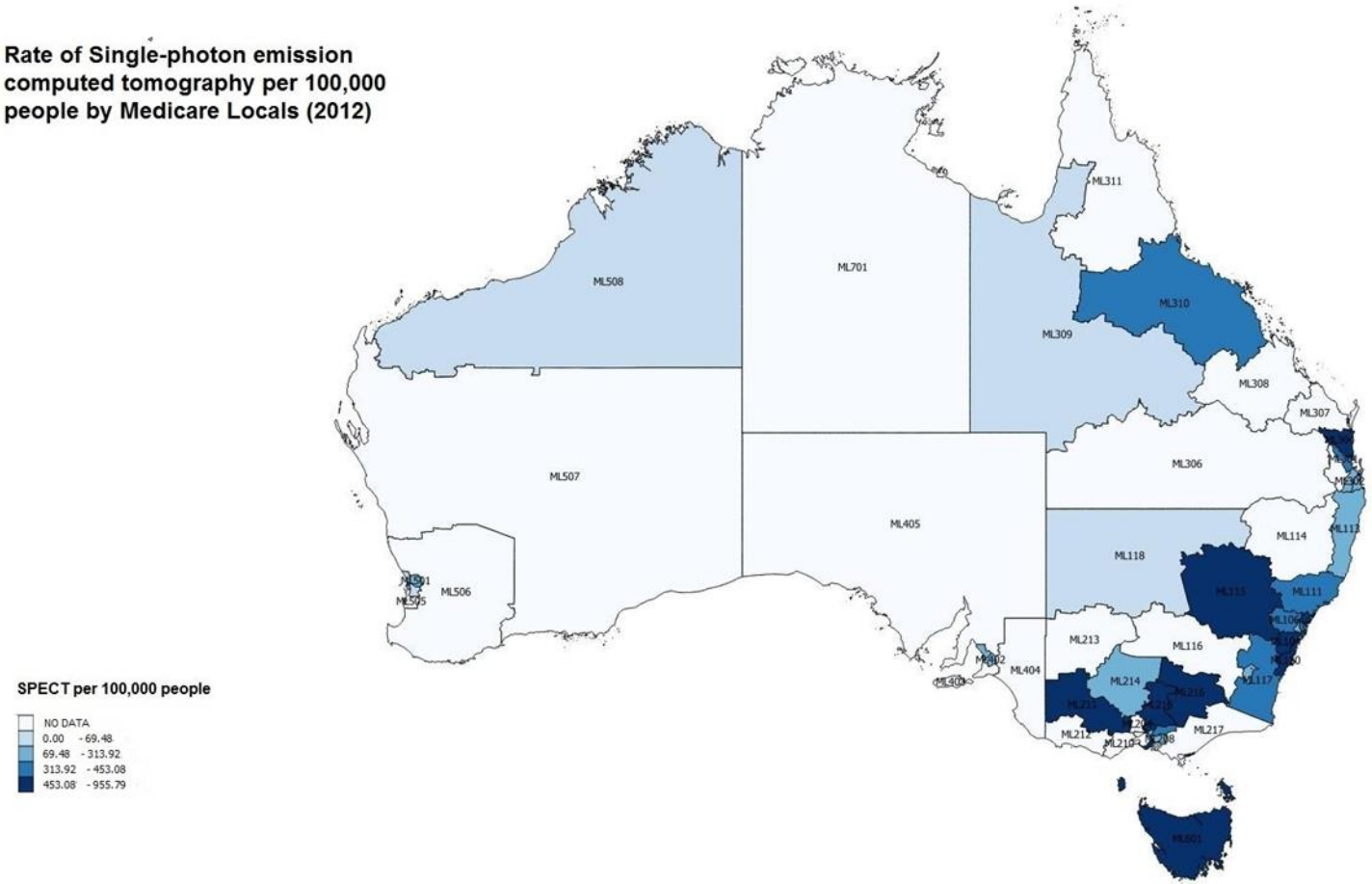


Figure 3- 6 Rate of single-photon emission computed tomography per 100,000 people (age-standardized) by Medicare Locals in Australia in 2012



Discussion

The results of this study demonstrated that there is an ongoing growth of cardiac imaging (transthoracic and transoesophageal echocardiography, stress echocardiography and single-photon emission computed tomography) per 100,000 people in Australia.

In addition, there is a substantial regional variation in the use of each of these modalities. As intense variations in practice are markers of potentially poor quality of care, this problem is worthy of attention.

Growth of cardiovascular imaging.

The sustainability of health systems has become an important point of interest in developed countries, due to increases in health expenditure, variations in practice, and proliferation in the provision of healthcare services.

Although the age-adjusted mortality rates for cardiovascular disease have shown a continuous decline due to advances in prevention and treatment, the extent to which cardiovascular imaging has helped this improvement in outcomes is unclear [96]. The latter is particularly a problem in relation to cardiac imaging [18, 89, 97]; in the United States Medicare population, cardiac imaging reimbursements grew from US$1.6 billion to US$5.1 billion between 2000 and 2006 [18]. At one stage, this increase was twice the average annual growth rate of all health care services [22], and in 2006, reimbursements paid for in-office cardiac imaging services amounted to 36% of the total Medicare part B payments [98].

Subsequently, a number of efforts have been made to control this growth, including changes in reimbursement models, the use of pre-authorisation using Radiology Benefits managers, patient education and the development of Appropriate Use Criteria for Cardiac imaging (AUC) [18, 99]. These efforts have been reasonably effective in managing the growth as we can identify a trend towards lower expenditure on cardiac imaging in the USA [25, 67], which started to appear around 10 years ago; however there have been ongoing concerns about the appropriate use of testing as physician behaviour has become reliant upon the control provided by third parties rather than self-management [6, 85].

In contrast to the reduction of cardiac imaging tests in the United States, the results of the present study show ongoing growth in Australia over the last decade, especially in

echocardiography rather than SPECT. This difference in trend among the modalities may reflect the performance of SPECT by specialised (nuclear medicine) physicians and radiology practices in Australia, which reduces access to this technology, in comparison to echocardiography, which is widely accessible.

Regional differences in imaging patterns.

In addition to its growth, the variation in the use of all cardiac imaging deserves special attention. The extremities of the range of all tests were broad; even after exclusion of outliers there remained a 15 to 47--fold variation in the use of these tests among regions.

Generally, greater regional use of testing was associated with the proportion with females, proximity to large cities, higher socioeconomic status and local concentration of physicians, but not with disease burden or deaths due to cardiovascular causes. Indeed, the regional medical workforce appears to be the strongest independent correlate of echocardiography use.

The most widely used cardiac imaging test was TTE, with a national rate of performance of 3,106 per 100,000 people per year. This rate of TTE utilization is less than the USA, where it averages 13,360 per 100,000 Medicare enrollees (range 4,000 to 34,000) [24], and greater than the UK average of 2,100 per 100,000 weighted population (range 120 to 4,200 per 100,000) [89, 100].

Some differences in the source of these data make comparisons of raw numbers difficult – Australian data captures outpatient echocardiograms, US data mainly derives from persons >65 years old, and UK data includes inpatient as well as outpatient echocardiograms. Interestingly, the rate of echocardiography has a 34 fold variation in the UK, and a 3.7 fold variation when the highest and lowest regions were excluded [89]. In this study, the numbers of trans-oesophageal and stress echocardiograms were too low to be meaningful in many areas, but the trends found were similar to the results of the use of TTE.

Although SPECT data were also not obtainable in all regions, different trends emerged. The association of SPECT use with non-metropolitan regions may reflect lower access to alternative testing (including stress echocardiography) in these regions. Importantly, SPECT was correlated with older population and cardiovascular death. This may be consistent with the more specific use of this test with coronary artery disease, rather than undifferentiated symptoms. The association with regional areas and older populations might imply that SPECT

has a greater potential for improving health outcomes than the current use of echocardiography [101].

This study supports the association of cardiac imaging with the location a patient seeks care [24]. Access to cardiac ultrasound is influenced by socioeconomic status, and strongly affected by the availability of doctors in the area, consistent with known associations between socioeconomic status, cardiovascular health and access to care [101, 102]. Indeed, the association of imaging with the number of doctors in the area was independent of distribution of patient gender, geography, burden of disease, and socioeconomic status.

Although potentially influenced by the fact that the data are restricted to outpatient echocardiograms (and therefore less morbid disease than would be observed in hospitalised patients), the use of all modalities of cardiac ultrasound was negatively associated with cardiovascular disease burden and mortality. Although echocardiography can be used to investigate nonspecific symptoms that may actually be caused by non-cardiovascular disease (e.g. dyspnoea finally attributed to lung disease), or to screening for cardiovascular disease that never actually eventuates (e.g. screening for LV dysfunction in chemotherapy), it appears unlikely that these and similar scenarios could fully explain this lack of association.

Geographical variations in health care practice may be explained by a number of reasons [88]. Repetition of previously performed tests may account for increased numbers in regions where there are numerous referral centres (e.g. urban Sydney, Melbourne and Adelaide). While financial incentives for performance of the test in a fee-for-service environment (generally USA, Australia) may explain some variations in practice, the similar regional variation in testing in the UK (where there is no financial incentive for additional screening) suggested that this explanation might be overstated [89]. More likely, variation in the rates of test utilisation reflect divergences in access to imaging and practice style [24].

Limitations.

This study has limitations that pertain to the use of administrative data. Access to details regarding disease burden, including differences among cardiovascular diseases and deaths due to cardiovascular causes, might help to distinguish tests performed for diagnosis, prognostic evaluation or follow-up. Particularly in the outpatient setting, the negative association between test use and proportion of cardiac diseases and deaths may reflect testing earlier in the disease course – which may indeed be appropriate. Patients with late stage disease are likely to have a

clear diagnosis and prognosis, so that any further investigation may add limited value. Likewise, access to additional patient-level information regarding the ordering physician, test indications, and proportion of repeat studies may improve understanding of appropriate use.

Finally, low utilisation in regional areas does not necessarily suggest that these patients are failing to have the imaging performed; a more likely scenario is these patients may attend a larger centre either by choice or because their disease requires them to visit a referral service.

Conclusion

The results of this study present evidence of growth and regional variation in the use of cardiac imaging in Australia, which is not completely accounted for by variations in demographics or disease burden.

The international profile of geographic variation in imaging use is surprisingly consistent, implying regional variations in practice style and/or access to imaging. Given the absence of association with variations in disease burden or mortality, it appears likely that reduction in this variation would not generate patient harm and could provide economic savings.

These observations suggest either under or over-utilization (or both), and implementation of an appropriate use process may generate convergence.

Chapter 4. Evidence status of the Appropriate Use Criteria

The research contained within this chapter has been published as[5]:

- Fonseca R, Negishi K and Marwick TH. What is the evidence status of Appropriate Use Criteria (AUC)? Insight from a matching exercise with the guidelines for echocardiography. *Intern Med J*. 2015;45:864-9.

Preface

In the previous chapter, growth in the use of cardiac imaging techniques within Australia, especially echocardiography, was demonstrated. Moreover, the findings indicate that the burden of disease does not explain the use of echocardiography. These results support the notion that there exists an inappropriate use of cardiac imaging in Australia.

As it was explained in Chapter 1, the AUC were designed to achieve appropriate use of cardiac imaging, but the challenges and potential risks of using the criteria in the Australian practice have not been assessed.

The following chapters, 4 to 7 (inclusive), aim to determine if the American AUC are in fact suitable to use in Australia in order to improve appropriate use of cardiac imaging particularly echocardiography, which, according to the findings in Chapter 3, needs an imminent intervention to ensure appropriate use among clinicians.

Several issues must first be addressed to determine if the AUC are suitable to use in Australia. Amongst these issues, assessing the scientific evidence status is crucial, due to the Australian practice of relying heavily upon the American guidelines for the management and diagnosis of cardiovascular diseases, including the use of echocardiography. Any discordance between those guidelines and the AUC would result in confusion with the attending clinician.

This chapter aims to determine the differences between the AUC and the published guidelines in order to understand possible challenges of the use of the appropriate use criteria.

Abstract

There is interest in adopting the American Appropriate Use Criteria (AUC) for transthoracic echocardiography to Australian practice. We matched 90 of 98 criteria of the AUC (53 "appropriate", 12 sometimes appropriate, 25 rarely appropriate) to the guidelines, but 8 criteria lacked any possible match. Among the matched criteria, 76 (82%) indications were concordant with the guidelines. A stronger basis in scientific evidence would be desirable to settle these discrepancies before Australian adoption of AUC.

Introduction

The growth of cardiac imaging has paralleled advances in the therapeutic options for cardiovascular disease and improvements in imaging technology [30]. However, the cost of this growth has necessitated a reduction in the heterogeneity of practice and provision of high-quality health services [22].

The Appropriate Use Criteria (AUC)[26] were developed in the USA to better align the expense and value of cardiac imaging [30]. However, some have expressed concern that the selection of AUC on the basis of expert opinion using a modified Delphi process risks the provision of unscientific guidance [64] because their evidence base is insufficiently robust [103].

In contrast, clinical guidelines have become a cornerstone of cardiology, optimising patient outcomes through facilitating clinical decision relating to the diagnosis, prevention or treatment of diseases or conditions [65]. These documents have classed recommendations according to the level of available scientific evidence [65], which has helped the calculation of risk and outcomes in clinical practice [104].

The adoption of AUC to Australian practice has been proposed as a means of containing the growth of imaging. Currently, Australian practice of cardiology relies heavily on the published American and European guidelines in regards to diagnosis and management of cardiovascular illnesses. Consequently, any conflicting messages, between the published guidelines and the appropriate use criteria, result in difficulties for the attending physician.

We hypothesised that the differing ways of formulating AUC and the guidelines might be responsible for discordances between these entities. Given the paucity of guidelines for transthoracic echocardiography in Australian literature, we were free to adopt both European and American guidelines so as to most exactly approximate the AUC.

In order to elucidate the relation between the class of recommendation and the appropriateness designation for transthoracic echocardiography, we sought to find a match for each item of ACCF/ASE/AHA Appropriate Use Criteria from published cardiology guidelines.

Methods

We searched for the matched items for each of the 98 criteria for transthoracic echocardiography in the 2011 AUC in the following manner: When a specific indication contained in the AUC lacked a match in the 2003 Guidelines for the Clinical Application of Echocardiography [65], other guidelines were used, principally in those related to management

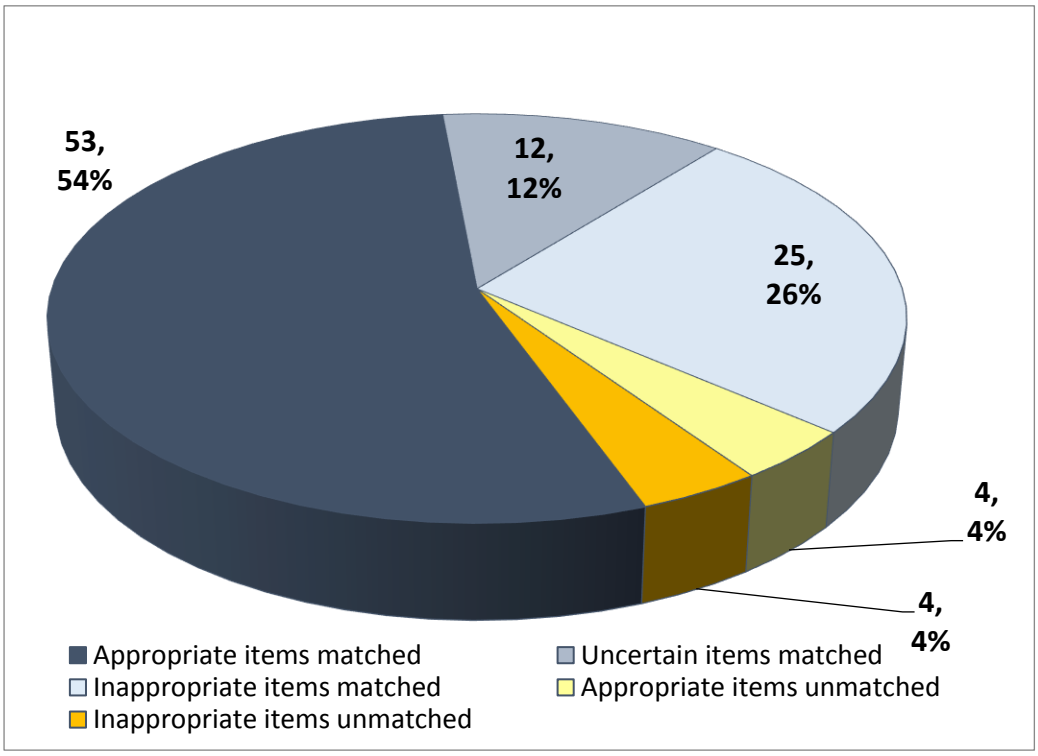
of heart failure [105], aortic disease [106], pulmonary hypertension [107], pulmonary embolism [108], cardiac devices [109], syncope [110], perioperative evaluation [111, 112], supraventricular arrhythmias [113], and valvular diseases [114].

We defined concordance if usually "appropriate" items (A) had Class I or IIa recommendation in the guidelines; "rarely appropriate" (RA) had Class III; or usually appropriate (UA) had Class IIb [30].

Results and discussion

Among the 98 AUC indications for the use of TTE, the majority (90 items, 92%) had a match in the different guidelines: 53 usually "appropriate", 12 sometimes appropriate, 25 rarely appropriate. Four of the usually "appropriate" indications and the same number of rarely appropriate indications did not have a counterpart in the guidelines (8% of the indications of the AUC for TTE) (Figure 4- 1).

Figure 4- 1 Appropriate Use Criteria indications with and without a counterpart in clinical guidelines



Blue: with matching. Yellow: no counterpart.

Among the 90 matched criteria, 76 (82%) indications were concordant with the guidelines, but there were 14 indications with discordance.

Four indications categorised as "Usually Appropriate" in the AUC had Class III recommendations in the guidelines. They were on syncope without cardiovascular signs and symptoms, routine surveillance of mild valvular stenosis and prosthetic valve without suspected valve dysfunction, and monitoring for rejection in cardiac transplant recipients.

Other two "appropriate" indications, on re-evaluation of known ascending aortic dilation or history of aortic dissection, where the concept was found in the guidelines, did not have any recommendation numbered and suggested limited use of TEE only to the root or those with Marfan syndrome (Table 4- 1).

There were five (20%) discordances in "inappropriate" ("rarely appropriate") indications (4 Class I and 1 Class IIa in guidelines), concerning the use of echocardiography in screening for heart disease, reevaluation of pulmonary hypertension with no change in clinical status, routine surveillance of moderate or severe valvular stenosis without a change in clinical status or cardiac exam and diagnosis of endocarditis or pulmonary embolism (Table 4- 2).

Table 4- 1 Discrepancies between Usually Appropriate indications and guidelines

AUC Table		Item No	AUC indication	Appropriate Use score		Guidelines*	Recommendation	Class
1	TTE for General Evaluation of Cardiac Structure and Function	9	Syncope when there are no other symptoms or signs of cardiovascular disease	A	7	1	Syncope in a patient for whom there is no clinical suspicion of heart disease.	III
3	TTE for Evaluation of Valvular Function	39	Routine surveillance (>3 y) of mild valvular stenosis without a change in clinical status or cardiac exam	A	7	1	Routine re-evaluation of asymptomatic adult patients with mild aortic stenosis having stable physical signs and normal LV size and function.	III
						1	Routine re-evaluation of asymptomatic patients with mild to moderate mitral stenosis and stable physical signs.	III
3	TTE for Evaluation of Valvular Function	49	Routine surveillance (≥3 y after valve implantation) of prosthetic valve if no known or suspected valve dysfunction	A	7	1	Routine re-evaluation of patients with valve replacements without suspicion of valvular dysfunction and unchanged clinical signs and symptoms.	III

6	TTE for Evaluation of Hypertension, HF or Cardiomyopathy	84	Monitoring for rejection in a cardiac transplant recipient	A	7	2	Recommendations for the Non-Invasive Monitoring of Acute Heart Transplant Rejection: Class III: 2. The use of echocardiography as an alternative to EMB for rejection monitoring is not recommended.	III
5	TTE for Evaluation of Aortic Disease	64	Re-evaluation of known ascending aortic dilation or history of aortic dissection to establish a baseline rate of expansion or when the rate of expansion is excessive.	A	9	3	Because TTE does accurately visualise the aortic root, its primary role as an imaging method for serial follow-up is in patients with aortic disease limited to the root, particularly those with Marfan syndrome. TEE is preferred.	N/A
5	TTE for Evaluation of Aortic Disease	65	Re-evaluation of known ascending aortic dilation or history of aortic dissection with a change in clinical status or cardiac exam or when findings may alter management or therapy	A	9	3	Because TTE does accurately visualise the aortic root, its primary role as an imaging method for serial follow-up is in patients with aortic disease limited to the root, particularly those with Marfan syndrome. TEE is preferred.	N/A

Guidelines*: 1, ACC/AHA Guidelines for the Clinical Application of Echocardiography[65]; 2, The International Society of Heart and Lung Transplantation Guidelines for the care of heart transplant recipients[112]; 3, 2010 ACCF/AHA/AATS Guidelines for the Diagnosis and Management of Patients With Thoracic Aortic Disease[106]; and 4. ACC/AHA 2006 Guidelines for the Management of Patients With Valvular Heart Disease [114]

Table 4- 2 Discrepancies between "Rarely Appropriate" indications ("Inappropriate") and guidelines

AUC Table		Item No	AUC indication	Appropriate Use score		Guidelines*	Recommendation	Class
1	TTE for General Evaluation of Cardiac Structure and Function	10	Initial evaluation of ventricular function (e.g., screening) with no symptoms or signs of cardiovascular disease	I	2	1	Patients with a family history of genetically transmitted cardiovascular disease.	I
						1	Patients with phenotypic features of Marfan syndrome or related connective tissue diseases.	I
						1	First-degree relatives (parents, siblings, children) of patients with unexplained dilated cardiomyopathy in whom no aetiology has been identified.	I
1	TTE for General Evaluation of Cardiac Structure and Function	16	Routine surveillance (<1 y) of known pulmonary hypertension without change in clinical status or cardiac exam	I	3	1	Follow-up of pulmonary artery pressures in patients with pulmonary hypertension to evaluate response to treatment.	I
3	TTE for Evaluation of Valvular Function	40	Routine surveillance (<1 y) of moderate or severe	I	3	1	Re-evaluation of asymptomatic patients with severe stenosis.	I

			valvular stenosis without a change in clinical status or cardiac exam					
3	TTE for Evaluation of Valvular Function	54	Transient bacteremia with a pathogen not typically associated with infective endocarditis and/or a documented nonendovascular source of infection	I	3	4	Transthoracic echocardiography to detect valvular vegetations with or without positive blood culture is recommended for the diagnosis of infective endocarditis.	I
2	TTE for Cardiovascular Evaluation in an Acute Setting	28	Suspected pulmonary embolism in order to establish diagnosis	I	2	1	Pulmonary emboli and suspected clots in the right atrium or ventricle or main pulmonary artery branches	IIa

Guidelines*: 1, ACC/AHA Guidelines for the Clinical Application of Echocardiography[65] and 4. ACC/AHA 2006 Guidelines for the Management of Patients With Valvular Heart Disease [114].

Additionally, there were three "Sometimes Appropriate" indications (Uncertain) (25%) with Class I recommendations in guidelines. These inconsistencies were related to evaluation of critically ill patients, re-evaluation of known HF with a change in clinical status or cardiac exam with a clear precipitating change in medication or diet and routine surveillance (<1 year) of adult congenital heart disease following incomplete or palliative repair with residual structural or hemodynamic abnormality without a change in clinical status or cardiac exam (Table 4- 3).

Discussion

In this matching exercise, we found significant inconsistencies with published guidelines in 15% of the 98 AUC for transthoracic echocardiography - 1% without indications in guidelines and 14% that contradict the guidelines.

The incorporation of clinical evidence into guidelines is an esential pillar of clinical practice, although the proliferation of guidelines, their variable quality and sometimes contradictory messages, make it a struggle for clinicians to incorporate these documents into their clinical practice [115]. Not all guidelines show a clear connection between evidence and recommendations[65]. Surprisingly – given the role of echocardiography as one of the most commonly used imaging techniques - the guidelines for the clinical application of echocardiography do not incorporate level of evidence[65, 105].

The Appropriate UC for echocardiography have been developed to help clinicians to choose testing more appropriately and to improve quality of care and standardisation of medical practice [26].

This study identified contradictions between AUC and guidelines, which are awkward and confusing, highlighting the shortcomings of a process whereby rating of appropriateness is based purely on expert opinion [64, 103]. A stronger evidence base is needed in order to settle these discrepancies in future updates of the AUC.

Table 4- 3 Discrepancies between Sometimes Appropriate Indications and guidelines

AUC Table		Item No	AUC indication	Appropriate Use score		Guidelines*	Recommendation	Class
2	TTE for Cardiovascular Evaluation in an Acute Setting	20	Assessment of volume status in a critically ill patient	U	5	1	The hemodynamically unstable patient.	I
6	TTE for Evaluation of Hypertension, HF or Cardiomyopathy	72	Re-evaluation of known HF (systolic or diastolic) with a change in clinical status or cardiac exam with a clear precipitating change in medication or diet	U	4	1	Re-evaluation of LV function in patients with established cardiomyopathy when there has been a documented change in clinical status or to guide medical therapy.	I
7	TTE for Adult Congenital Heart Disease	97	Routine surveillance (<1 y) of adult congenital heart disease following incomplete or palliative repair + with residual	U	5	1	Periodic echocardiography in patients with surgically repaired (or palliated) congenital heart disease with the following: change in clinical condition or clinical suspicion of residual defects,	I

			structural or hemodynamic abnormality + without a change in clinical status or cardiac exam				obstruction of conduits and baffles, or LV or RV function that must be followed, or when there is a possibility of hemodynamic progression or a history of pulmonary hypertension	

Guidelines*: 1, ACC/AHA Guidelines for the Clinical Application of Echocardiography [65]

Given the number of guidelines evaluating the use of echocardiography in different diseases and scenarios, our study results do not show all the possible relationships between the AUC criteria and the published guidelines.

We showed the main differences between American AUC and American and European guidelines, but it is likely that more discrepancies would become apparent if additional guidelines were compared. The practice of echocardiography in Australia parallels, and is informed by, the practices of the American and European cardiology societies.

It is imperative to develop a greater degree of internal consistency between guidelines and the AUC to facilitate widespread adoption of the AUC.

Conclusion

The appropriate use criteria for transthoracic echocardiography are not consistent with the guidelines for the use of echocardiography.

The potential incorporation of an AUC process into Australian practice might still improve patient outcomes, reduce variation and contain costs, but should be informed by these limitations.

Chapter 5. Impact of Appropriate Use Criteria on clinicians' behaviour

The research contained within this chapter has been published as[6]:

- Fonseca R, Negishi K, Otahal P and Marwick TH. Temporal changes in appropriateness of cardiac imaging. *J Am Coll Cardiol*. 2015;65:763-73.

Preface

In the previous chapter, a comparison between the AUC for echocardiography and clinical guidelines was assessed. The results showed that the AUC are not entirely concordant with the published cardiovascular guidelines, which if adopted as is would bring issues in the Australian clinical practice.

In order to continue with the evaluation of the AUC to determine their suitability for use in Australia, this chapter aims to evaluate if the AUC are a tool that influences clinicians' ordering behaviour. This will be done by assessing the changes of proportion of "appropriate" and inappropriate requests since the AUC were first launched.

This chapter documents a systematic review of the published literature determining changes in the proportion of appropriate and inappropriate tests over time as an indicator of impact on clinicians' requesting behaviour.

Abstract

Appropriate use criteria (AUC) for cardiac imaging have been available for more than ten years. The extent to which improvements in appropriate use have been identified and reported are undefined. This study systematically reviewed published evidence to identify whether the promulgation of the AUC has led to an improvement in the proportion of appropriate cardiac imaging requests.

Methods:

Electronic databases were systematically searched for English-language papers related to AUC and cardiovascular imaging. We found 59 reports involving 103,567 tests that were published between 2000 and 2012. The rate of appropriate testing over time was analysed in a meta-regression.

Results:

New AUC were associated with apparent improvement of appropriateness for transthoracic echocardiography (TTE) (80% [95% confidence interval (CI): 0.75 to 0.84] in 2007 to 85% [95% CI: 0.81 to 0.89] in 2011), trans-oesophageal echocardiography (TEE) (89% [95% CI: 0.81 to 0.94] in 2007 vs 95% [95% CI: 0.93 to 0.96] in 2011), and computed tomography angiography (CTA) (37% [95% CI: 0.21 to 0.55] in 2006 vs 55% [95% CI: 0.44 to 0.65] in 2010), but not for stress echocardiography (SE) (53% [95% CI: 0.45 to 0.61] in 2008 vs 52% [95% CI: 0.42 to 0.61] in 2011) or single-photon emission computed tomography (SPECT) (72% [95% CI: 0.66 to 0.77] vs 68% [95% CI: 0.60 to0.74] in 2005 and 2009 respectively). Although there were no correlations between the proportion of appropriate TTEs and published year (p=0.36) for 2007 AUC, there was a positive correlation between proportion of appropriateness and the year of publication (p=0.01) for 2011 AUC. There was a significant decrease in the proportion of appropriateness over time using the 2007 TEE AUC (p=0.03) and 2006 CTA AUC (p=0.02). There were no meaningful associations between appropriateness and publication year for stress echocardiography, CTA or SPECT.

Conclusion: Rates of reported appropriate use in imaging show improvements for TTE and CTA but not for stress imaging and TEE. The observed reductions in imaging studies are not matched by reported rates of appropriate use.

Introduction

The Appropriate Use Criteria (AUC) were launched to reduce heterogeneity of practice and to improve health service quality. Subsequently, AUC have also been considered as a method of controlling resource utilisation and medical expenditures [31]. This is specifically true in cardiac imaging, where at one stage; the growth in costs was double the average annual increase of all services [99]. Cardiology imaging reimbursements increased from US $1.6 billion in 2000 to US $5.1 billion in 2006 [18]. However, after 2009, the volume of cardiac imaging has shown a decreasing trend [25, 67, 116].

The initial AUC for single photon emission computed tomography (SPECT) were followed by AUC for cardiac magnetic resonance (CMR) and cardiac computed tomography (CCT), transthoracic echocardiography (TTE), transesophageal echocardiography (TEE), and stress echocardiography (SE) [26, 32, 33, 117-119].

It is unclear whether the publication of AUC was the reason for the reduction of cardiac imaging. Evaluations after educational campaigns have shown heterogeneous responses [120-122]. When changes have been reported, some papers have shown a decline in the number of requests[116, 120, 121], and most have shown an improvement in the proportion of tests coded as being appropriate [122].

The goal of the current study was to demonstrate the impact if any, of AUC on the ordering behaviour of clinicians by examining reported rates of appropriateness over time.

Methods

Search strategy.

We adhered to the protocol specified in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement for reporting systematic reviews [123].

Two reviewers (R.F and K.N) conducted a literature search of five online databases (Pubmed / Medline, Embase, Web of Science, Scopus , and Cochrane) for published studies that estimated the proportion of appropriate tests using the AUC for cardiac imaging including TTE, TEE, SE, SPECT, CMR and CCT for all years from 2005 to 2014.

Search keywords included the terms ["Cardiac" OR "Cardiovascular"] AND ["imaging" OR "echocardiography" OR "echocardiogram" OR "nuclear cardiology" OR "SPECT" OR "single-photon emission computed tomography" OR "single photon emission computed

tomography" OR "Positron emission tomography" OR "CCT" OR "cardiac computed tomography" OR "computed tomography" OR "CT" OR "CMR" OR "cardiac magnetic resonance" OR "magnetic resonance"] AND ["appropriate* use criteria" OR "AUC"]. Papers were limited to those published in English. References to publications and relevant articles were also searched for further reports.

Inclusion criteria.

Publications in peer-reviewed, English language journals evaluating the AUC in echocardiography (TTE, TEE, and SE), SPECT, CMR and CTA were included in this systematic review if they reported the following:

- The Appropriate Use Criteria edition and type used;
- The year of collection of the data;
- The sample size of tests evaluated;
- The proportion of appropriate studies and the proportion of uncertain ("maybe appropriate") or inappropriate ("rarely appropriate") tests;
- The proportion of classified and/or unclassified tests.

No restrictions were applied to the types of patients, the report's country of origin or type of institution where the AUC were evaluated.

Outcomes.

The primary outcome of this systematic review was to calculate the proportion of appropriate tests of the total sample (defined by the proportion of appropriate tests in the total sample of each study). The secondary outcome was to determine the proportion of appropriate tests of classifiable imaging studies. For both, our goal was to establish the relation between appropriateness with the median year of acquisition of the data and year of publication of the manuscript. We also sought to assess the trends in classified imaging studies over time for each of the AUC editions.

Data extraction.

Data were extracted by one review author (R.F.) and checked by a second reviewer (K.N.). Discrepancies between reviewers were resolved by consensus, or, if necessary, by a third author (T.H.M.).

Information on publication year, average of enrolment year, sample size, proportion of appropriate tests, gender, mean age, proportion of inpatient population and speciality of doctors who requested the tests was extracted independently from every eligible report.

The goal of this extraction was to determine whether there was any relationship between appropriateness and these factors. If a paper reported the effect of intervention in the same population but at two points in time, both data points were used in two different analyses.

Because there were two editions of the AUC for each cardiac imaging test at the time of data extraction, the analysis was performed in 10 groups: 2007 and 2011 TTE, 2007 and 2011 TEE, 2008 and 2011 SE, 2005 and 2009 SPECT, and 2006 and 2010 CCA.

Statistical analysis.

Meta-analysis was performed using a logit transformation to calculate the weighted summary proportion under a random-effects model (DerSimonian-Laird estimator). We assumed that effect sizes differed between studies due to variances in the characteristics of participants, dissimilarities between hospitals and regions, or protocols for using cardiac imaging in diverse scenarios, among others. A random effect model was used because the variation in observed effects were not only due to sampling error.

Ten pooled analyses were conducted separately: one for each cardiac imaging and edition of AUC used (TTE 2007, TTE 2011, TEE 2007, TEE 2011, SE 2008, SE 2011, SPECT 2005, SPECT 2009, CTA 2006 and CTA 2010).

Reports were included in more than one analysis if the study evaluated more than one cardiac imaging technique or a different edition of AUC. Heterogeneity between the included studies was assessed using Cochrane's Q (reported with a chi-square value and p-value) and was quantified with the I^2 statistic.

Possible sources of heterogeneity were investigated further by meta-regression analysis. The characteristics selected for this study included age, gender, hospitalisation status, specialisation of the physician who referred for testing, year of publication and the average year of enrolment of data.

Meta-regressions were performed on the (non-linear) logit scale; to show the effect of each study characteristic, we back-transformed the model coefficient at the mean of each characteristic. For those characteristics that were proportion data (inpatients, male, specialists)

we have shown the coefficient per 0.1 unit (or 10%) increase in proportion rather than the standard per 1-unit increase. Publication bias was examined by plotting a funnel plot and was quantified by Egger's test.

All statistical analyses were performed with R software version 3.1.0 with the following packages: "meta", "metaphor", and "boot"[79].

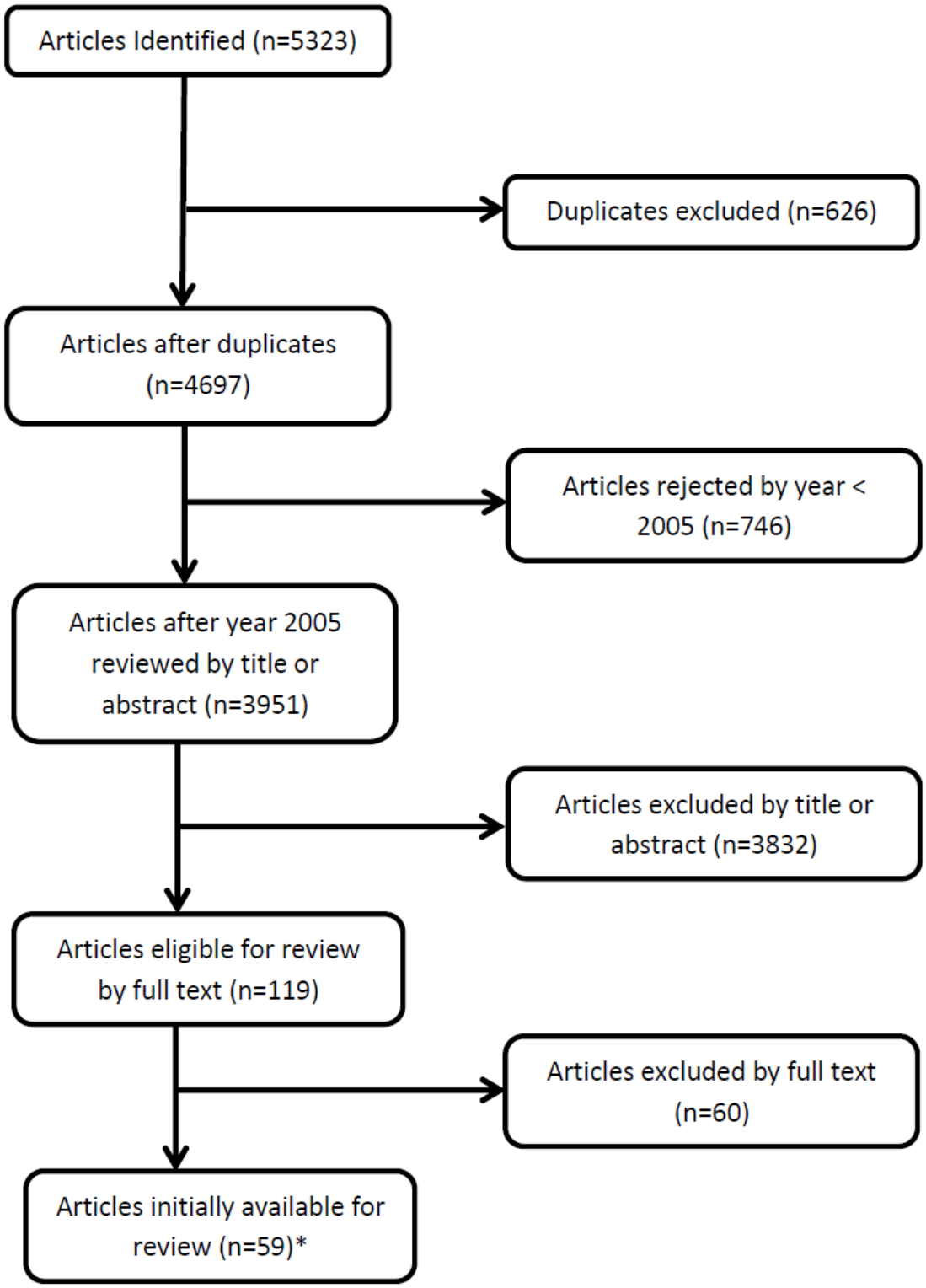
Results

Literature search process.

The initial search of the five online databases used in this work identified a total of 5,323 original papers (Figure 5- 1). Exclusion of 5,264 after reviewing of the title, abstract or both, left 59 possible articles suitable for the present systematic review.

Some studies were used more than once if they had data for different cardiac imaging or the same imaging but different AUC edition. Of the studies included, 15 studies were used for the TTE 2007 AUC analysis; 10 studies for TTE 2011 AUC, 5 studies for TEE 2007 AUC, 3 studies for TEE 2011 AUC, 6 for SE 2008 AUC, 8 for SE 2011 AUC, 11 for SPECT 2005 AUC, 11 for SPECT 2009 AUC, 9 for CTA 2006 AUC, and 7 for CTA 2010 AUC analysis.

Figure 5- 1 . Schematic diagram of literature search and selection procedure for articles included in the systematic review



Characteristics of original reports.

The majority of reports (53%) were retrospective in design (Appendix Table 4). An individual different to the ordering physician did the appropriateness scoring by using medical records, data and requests of tests in nearly 90% of studies. Among these studies in which reviewers scored appropriateness, 58% had the score reviewed only for unclassified tests or disagreement between reviewers.

Physicians evaluated appropriateness in 51% of the studies, nurses or sonographers were reviewers in 15%, and the occupation of the reviewer was undefined in 34%. The weighted average of appropriate tests for physicians was 40% of the total sample, compared to 65% for nurses and 71% for sonographers.

In the vast majority of the studies, there were no reports about the agreement between reviewers. In those papers in which reviewer agreement was reported, the agreement (kappa) between physician reviewers varied between 0.31 and 0.84. This range exceeded that for nurses (0.56 to 0.74) and sonographers (0.67 and 0.84).

The observations were based on the request at the point-of-service in 86%. Most (73%) were performed in an academic setting, with 13.6% in a community setting and 13.6% in both environments.

Table 5- 1 and Table 5- 2 present the studies included for each of the analyses.

Table 5- 1 Overview of studies included for the TTE and TEE analysis

Study	Test	AUC Edition	Publ Year	Enrolment year	n	App tests	App (%)	Classified Studies	Previous echo (%)	Inpatient (%)	Age (y)	Women	Cardiac Specialist %
Ward[59]	TTE*	2007	2008	2007	1553	1228	0.79	1385	0.36	0.48	58.8±16.9	0.53	0.48
Willens[60]	TTE	2007	2009	2008	625	481	0.77	526	_	0.17	_	_	0.22
Dharmarajan[46]	TTE	2007	2009	2003	58	51	0.88	58	_	_	29.0±6.0	1.00	0.19
Kirkpatrick[49]	TTE	2007	2009	2007	368	206	0.56	237	0.78	0.00	55.0±17.0	0.51	0.61
Martin[51]	TTE	2007	2009	2008	274	237	0.86	268	_	1.00	_	0.50	0.38
Bhave [45]	TTE	2007	2010	2009	258	199	0.77	221	0.35	_	59.0±18.0	0.53	0.55
Rao[56]	TTE	2007	2010	2008	772	533	0.69	716	_	0.00	_	_	1.00
Aggarwal[38]	TTE	2007	2010	2007	329	278	0.84	299	_	0.44	63.0±15.0	0.42	0.57
Gathak[47]	TTE	2007	2011	2009	431	364	0.84	394	_	_	_	_	_
Rahimi1[55]	TTE	2007	2011	2000	177	143	0.81	164	0.37	0.00	53.0±17.0	0.27	0.37
Rahimi2[55]	TTE	2007	2011	2008	348	251	0.72	296	0.54	0.00	58.0±17.0	0.48	0.37
Parikh[53]	TTE	2007	2012	2010	384	333	0.87	336	0.31	0.68	64.0±16.0	0.45	0.51
Bhatia [42]	TTE	2007	2012	2011	450	288	0.64	347	0.69	0.33	70.6±14.7	0.50	0.38
Alqarqaz[39]	TTE	2007	2012	2009	170	131	0.77	147	_	_	_	_	0.40

Silverman[58]	TTE	2007	2012	2009	485	442	0.91	485	_	_	_	0.50	_
Bailey[40]	TTE	2007	2013	2008	1080	933	0.86	945	_	1.00	71.2±15.0	_	0.10
Willens[61]	TTE	2011	2011	2008	625	479	0.77	617	_	0.17	_	_	0.22
Parikh[53]	TTE	2011	2012	2010	384	354	0.92	363	0.31	0.68	64.0±16.0	0.45	0.51
Bhatia1[42]	TTE	2011	2012	2011	450	313	0.69	441	0.69	0.33	70.6±14.7	0.50	0.38
Patil[54]	TTE	2011	2012	2010	1820	1493	0.82	1812	_	0.47	_	_	_
Alqarqaz[39]	TTE	2011	2012	2009	170	131	0.77	170	_	_	_	_	0.40
Ballo[41]	TTE	2011	2012	2010	931	739	0.79	920	_	_	72.8±14.4	0.46	0.49
Mansour[50]	TTE	2011	2012	2007	1553	1253	0.81	1525	_	0.49	59.0±17.0	0.52	0.50
Bailey [40]	TTE	2011	2013	2008	1080	1042	0.96	1080	_	1.00	71.2±15.0	_	0.10
Matulevicius[35]	TTE	2011	2013	2011	535	491	0.92	535	_	0.57	_	0.59	0.31
Bhatia2[120]	TTE	2011	2013	2011	1318	1105	0.84	1312	_	_	63.0	0.46	_
Bhatia3[120]	TTE	2011	2013	2012	345	312	0.90	337	_	_	61.0	0.42	_
Rao[57]	TEE †	2007	2009	2006	1235	1156	0.94	1235	_	_	61.0	_	_

Aggarwal[38]	TEE	2007	2010	2007	200	191	0.95	194	_	0.50	63.0±15.0	0.42	0.57
Ogbara[52]	TEE	2007	2011	2011	389	321	0.82	389	_	_	_	_	_
Bhatia[43]	TEE	2007	2012	2011	202	156	0.77	166	_	0.76	63.0±14.0	0.20	0.73
Mansour[50]	TEE	2007	2012	2007	405	358	0.88	368	_	0.72	59.0±17.0	0.52	0.50
Bhatia[43]	TEE	2011	2012	2011	202	190	0.94	199	_	0.76	63.0±14.0	0.20	0.73
Grewal[48]	TEE	2011	2012	2008	671	639	0.95	659	_	_	66.0±13.0	0.33	_
Mansour [50]	TEE	2011	2012	2007	405	382	0.94	404	_	0.72	59.0±17.0	0.52	0.50

*TTE: Transthoracic Echocardiography

†TEE: Trans-oesophageal Echocardiography

Table 5- 2 Overview of studies included for the SE, SPECT and CCT analysis

Study	Imaging test	AUC Edition	Publ Year	Enrolment year	n	App tests	App (%)	Classified Studies	Previous echo (%)	Inpatient (%)	Age (y)	Women	Cardiac Specialist%
McCully[124]	SE*	2008	2009	2005	298	159	0.53	241	_	_	66.0±13.0	0.48	_
Mansour[125]	SE	2008	2010	2008	289	180	0.62	253	_	0.07	59.0±18.0	0.49	0.45
Bhatia[44]	SE	2008	2013	2011	252	104	0.41	126	0.15	_	58.1±12.2	0.42	0.50
Willens[121]	SE	2008	2013	2008	209	104	0.50	189	_	0.00	56.1±13.8	0.53	0.52
Lin[126]	SE	2008	2013	2010	111	50	0.45	92	_	_	51.4	0.54	_
Schmitz[63]	SE	2008	2013	2010	300	194	0.65	226	_	_	_	_	
Mansour[50]	SE	2011	2012	2008	289	165	0.57	281	_	0.07	59.0±17.0	0.52	0.50
Cortigiani[127]	SE	2011	2012	2003	1552	984	0.63	1552	_	0.00	_	_	_
Bhatia[44]	SE	2011	2013	2011	252	105	0.42	221	0.15	_	58.1±12.2	0.42	0.50
Bhattacharyya [62]	SE	2011	2013	2011	100	49	0.49	100	_	_	_	_	_
Willens1[121]	SE	2011	2013	2008	209	100	0.48	207	_	0.00	56.1±13.8	0.53	0.52
Willens2[121]	SE	2011	2013	2011	209	82	0.39	200	_	0.00	56.3±14.7	0.53	0.53
Willens3[121]	SE	2011	2013	2001	111	48	0.43	107	_	0.00	57.7±13.3	0.50	1.00
Schmitz[63]	SE	2011	2013	2010	300	300	1.00	300	_	_	_	_	
Gibbons1[128]	SPECT	2005	2008	2005	284	182	0.64	253	_	_	67.0±11.0	0.37	_

Mehta[129]	SPECT	2005	2008	2006	1209	940	0.78	1173	_	_	_	0.55	0.69
Hendel [130]	SPECT	2005	2010	2007	6351	4192	0.66	5906	_	_	65.7±11.8	0.41	0.75
Gibbons2[131]	SPECT	2005	2010	2006	284	188	0.66	241	_	_	68.0±11.0	0.33	_
Carryer[132]	SPECT	2005	2010	2005	281	179	0.64	250	_	_	67.0±11.0	0.37	_
Gupta[133]	SPECT	2005	2011	2009	314	263	0.84	314	_	_	62.0±14.0	0.48	0.62
Gibbons3[122]	SPECT	2005	2011	2008	273	164	0.60	232	_	_	65.0±13.0	0.33	_
Gholamrezanezhad[134]	SPECT	2005	2011	2009	291	211	0.72	279	_	_	55.3±10.3	0.57	_
Druz [135]	SPECT	2005	2011	2007	585	370	0.63	570	_	0.48	63.5±13.1	0.45	0.44
Soine1[136]	SPECT	2005	2012	2007	1377	950	0.69	1377	_	_	58.4±13.4	0.52	_
Soine2[136]	SPECT	2005	2012	2007	1445	1286	0.89	1445	_	_	60.8±10.6	0.09	_
Carryer[132]	SPECT	2009	2010	2005	281	168	0.60	281	_	_	67.0±11.0	0.37	_
Koh[137]	SPECT	2009	2011	2009	1623	1331	0.82	1574	_	_	61.0±11.0	0.39	0.93
Gholamrezanezhad[134]	SPECT	2009	2011	2009	291	219	0.75	283	_	_	55.3±10.3	0.57	_
Nelson1[138]	SPECT	2009	2012	2009	150	101	0.67	148	_	0.12	61.0±10.0	0.01	_
Nelson2[138]	SPECT	2009	2012	2009	150	111	0.74	150	_	_	65.0±12.0	0.43	0.47
Koh[139]	SPECT	2009	2012	2009	176	106	0.60	176	_	_	61.0±11.0	0.41	_
Lin[126]	SPECT	2009	2013	2010	338	178	0.53	312	_	_	57.3	0.34	1.00
Winchester[140]	SPECT	2009	2013	2011	332	259	0.78	328	_	_	_	0.04	_

Doukky[141]	SPECT	2009	2013	2009	1511	779	0.52	1491	_	_	59.0±13.0	0.43	_
Moralidis[142]	SPECT	2009	2013	2011	3032	2208	0.73	3008	_	_	66.0±11.0	0.41	_
Aldweib[143]	SPECT	2009	2013	2006	1199	740	0.62	1194	0.36	_	63.8±12.5	0.44	_
Ayyad 1[144]	CCT ‡	2006	2009	2006	763	530	0.69	715	_	_	57.2±13.6	0.35	_
Ayyad 2[144]	CCT	2006	2009	2007	646	507	0.78	623	_	_	58.1±13.3	0.35	_
Miller[145]	CCT	2006	2010	2007	251	69	0.27	136	_	_	_	_	_
Murphy[146]	CCT	2006	2010	2008	267	126	0.47	189	0.69	_	56.2±14.0	0.36	0.82
El Sibai [147]	CCT	2006	2011	2009	100	8	0.08	100	_	0.19	53.0±13.0	0.17	0.77
Chinnaiyan[148]	CCT	2006	2012	2009	25387	5053	0.20	12853	_	0.27	57.0	0.46	0.21
Rich[149]	CCT	2006	2012	2011	1216	503	0.41	1069	_	0.31	57.5±15.7	0.47	_
Mazimba[150]	CCT	2006	2012	2007	243	36	0.15	243	_	_	59.2±12.3	0.55	_
Wasfy[151]	CCT	2006	2012	2008	267	119	0.45	189	0.51	_	56.2±14.0	0.36	_
El Sibai [147]	CCT	2010	2011	2009	100	38	0.38	100	_	0.19	53.0±13.0	0.17	0.77
Chinnaiyan[148]	CCT	2010	2012	2009	25387	18266	0.72	22442	_	0.27	57.0	0.46	0.21
Rich[149]	CCT	2010	2012	2011	1216	863	0.71	1159	_	0.31	57.5±15.7	0.47	_
Mazimba[150]	CCT	2010	2012	2007	243	119	0.49	243	_	_	59.2±12.3	0.55	_
Wasfy[151]	CCT	2010	2012	2008	267	157	0.59	231	0.51	_	56.2±14	0.36	_
Lin[126]	CCT	2010	2013	2010	23	13	0.56	18	_	_	50.3	0.39	_
Cullen [152]	CCT	2010	2013	2007	251	85	0.34	212	_	_	_	_	_

*SE: Stress Echocardiography

†SPECT: Single-photon emission computed tomography

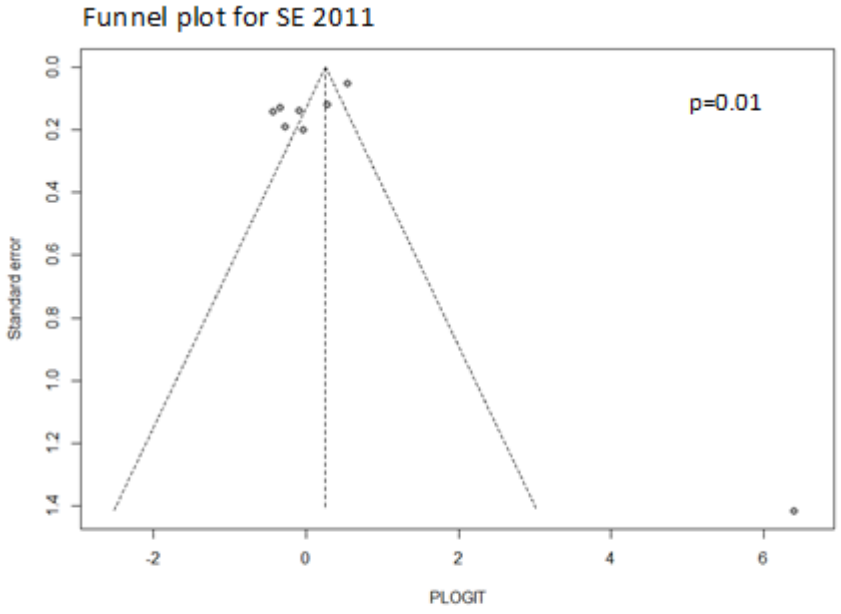
‡CCT: Cardiac Computed Tomography

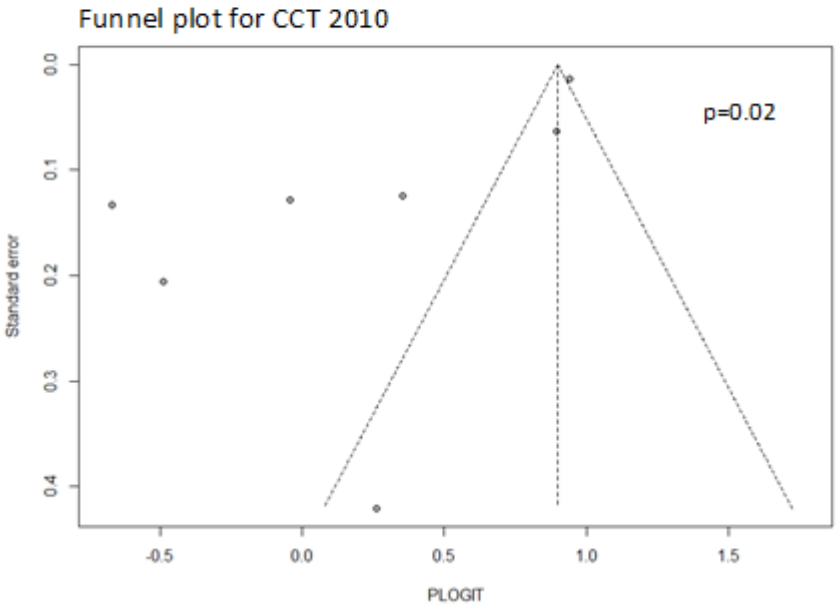
Publication bias.

Funnel plots were made for each of the pooled analyses (Figure 5- 2). There was evidence for publication bias for stress echocardiography 2011 (p=0.01) and CTA 2010 (p=0.02).

However, bias was not identified for TTE 2007 (p=0.24), TTE 2011 (p=0.10), TEE 2007 (p=0.99), TEE 2011 (p=0.39), SE 2008 (p=0.65), SPECT 2005 (p=0.35), SPECT 2009 (p=0.93), CTA 2006 (p=0.08).

Figure 5- 2 Publication bias for Stress echocardiography (2011 AUC edition) and Cardiac computed tomography (2010 AUC edition)



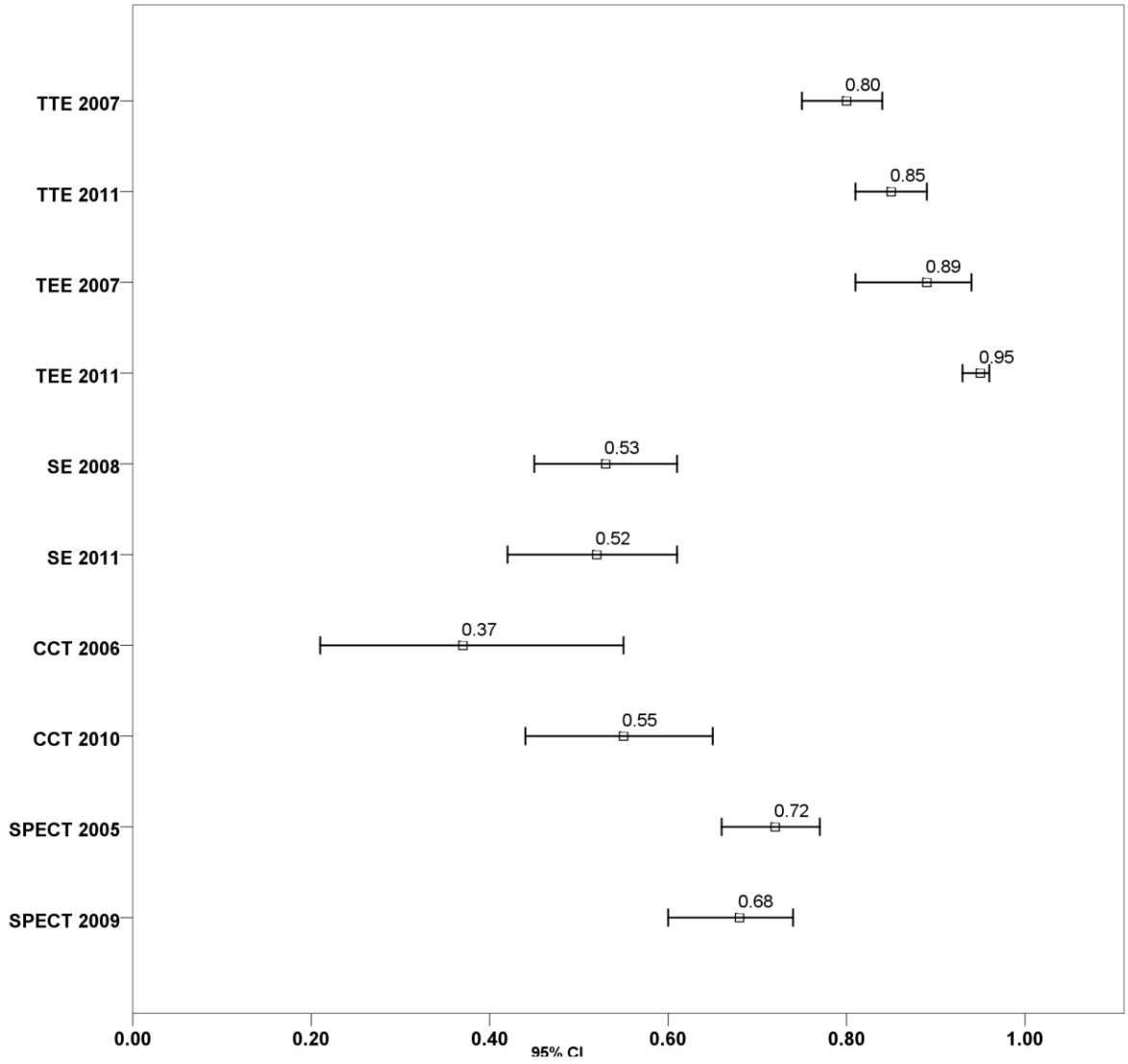


Temporal changes in appropriate use.

The overall findings of this study are that new versions of AUC were associated with apparent improvement of appropriateness for transthoracic (TTE) and transesophageal echocardiography (TEE), and computed tomography angiography (CTA), but not for stress echocardiography (SE) or single photon emission computed tomography (SPECT) (Figure 5-3).

The only modality showing a correlation between the proportion of appropriate testing and the published year was for TTE using the 2011 guidelines.

Figure 5- 3 Reported Appropriate Use in Papers Applying Different Versions of Appropriate Use Criteria



Each line represents a summary of overall estimates calculated from random effect models of the proportion of appropriate tests. Individual Forest plots are provided in the online figures. TTE: Transthoracic echocardiography; TTE: Trans-oesophageal echocardiography, SE: Stress echocardiography, CCT: Cardiac computed tomography, SPECT: Single-photon emission computed tomography.

(Y axis: AUC edition, X axis: Overall estimates of appropriate use (95% CI)).

Transthoracic echocardiography (2007 AUC edition).

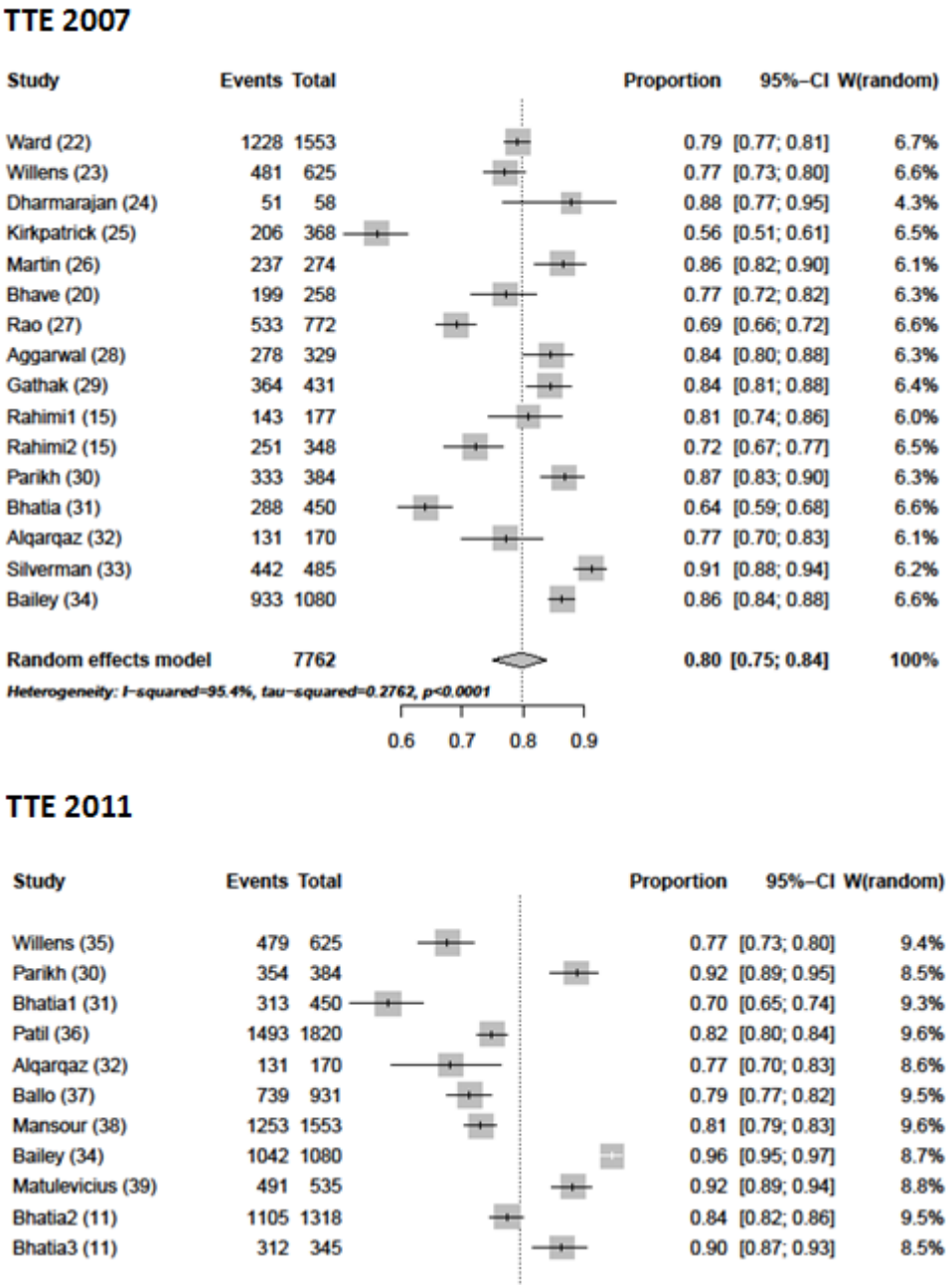
Of the 15 studies in which the of AUC for TTE was evaluated using the 2007 edition, one study[55] was presented twice because it analysed two different samples for TTE at two different of points of time (Table 5- 1, Figure 5- 4). Thus, there are 16 different rows (n=7,762 TTE) analysed for 15 different studies published between 2008 and 2013, with a range of enrolment between 2000 and 2011.

The pooled proportion of appropriateness of the total sample was 0.80 [95% CI, 0.75 to 0.84], with substantial heterogeneity among the estimates [I^2= 95.4%, p<0.0001]. The weighted average of appropriate tests among classifiable studies was 0.91 [95% CI: 0.87 to 0.93] (Table 5- 3).

We explored heterogeneity using the following factors: publication year, gender, proportion of inpatients and proportion ordered by cardiologists. In univariable meta-regression, there were no significant associations of the proportion of appropriate tests with publication year (p=0.36) (Table 5- 4).

In addition, there were no significant relations between appropriateness and gender, type of patient, or cardiologist who ordered the test. However, there was a positive association between the proportion of appropriate testing and the proportion of inpatients (p=0.0006) (Table 5- 3). The proportion of classifiable studies did not show a significant improvement over time (Table 5- 3).

Figure 5- 4 Appropriateness of Transthoracic Echocardiography using the 2007 (up) and 2011 (down) AUC edition



Studies were ordered by increasing publication year. The overall estimate was calculated from random effect model.

Table 5- 3 Proportion of appropriate tests and their association with the publication year in the total sample (first columns) and among classifiable studies (2nd group of columns). The 3rd group of columns examines the association between the proportion of classifiable tests and publication years.

	Total sample			**Classified studies**			**% Classified studies of Total sample**		
	%Appropriate	Pub year*	p	%Appropriate	Pub year*	p	%Classified	Pub year*	p
TTE 2007	0.80 [0.75;0.84]	0.09	0.36	0.91 [0.87;0.93]	0.26	0.06	0.89 [0.85;0.92]	-0.02	0.87
TTE 2011	0.85 [0.81;0.89]	0.73	0.01	0.87 [0.83;0.90]	0.69	0.07	0.99 [0.98;0.99]	0.67	0.24
TEE 2007	0.89 [0.81;0.94]	-0.43	0.03	0.95 [0.89;0.97]	-0.01	0.98	0.97 [0.91;0.99]	-1.54	0.03
TEE 2011	0.95 [0.93;0.96]	–	–	0.96 [0.94;0.97]	–	–	0.99 [0.97;0.99]	–	–
SE 2008	0.53 [0.45;0.61]	-0.07	0.46	0.71 [0.60;0.80]	0.05	0.80	0.80 [0.67;0.89]	-0.12	0.57
SE 2011	0.52 [0.42;0.61]	-0.04	0.97	0.53 [0.44;0.61]	-1.42	0.35	0.98 [0.95;0.99]	-1.42	0.35
SPECT 2005	0.72 [0.66;0.77]	0.10	0.40	0.76 [0.71;0.80]	0.04	0.70	0.95 [0.92;0.97]	0.84	0.05
SPECT 2009	0.68 [0.60;0.74]	-0.12	0.42	0.69 [0.71;0.76]	-0.13	0.42	0.99 [0.97;0.99]	-0.01	0.99
CCT 2006	0.37 [0.21;0.55]	-0.59	0.02	0.48 [0.35;0.62]	-0.57	0.05	0.87 [0.74;0.94]	-0.09	0.87
CCT 2010	0.55 [0.44;0.65]	-0.06	0.89	0.61 [0.47;0.74]	0.16	0.76	0.90 [0.85;0.94]	-1.41	0.03

*Regression coefficient (slope). Values in brackets are 95% confidence intervals. Abbreviations as in Tables 1 and 2.

Table 5- 4 Meta- regression of the associations of the proportion of appropriate tests.

	Publication year*	**p**	**Male***	**p**	**Cardiologists***	**p**	**Inpatient***	**p**	**Age***	**p**
TTE 2007	0.09	0.36	-0.62	0.64	-1.03	0.10	1.10	<0.01	-0.01	0.67
TTE 2011	0.73	0.01	-0.02	1.00	-2.59	0.22	3.04	<0.01	0.00	0.99
TEE 2007	-0.43	0.03	-3.42	0.42	-4.81	0.42	-6.34	0.01	-0.01	0.97
TEE 2011	–	–	0.16	0.88	0.16	0.88	0.16	0.88	0.03	0.51
SE 2008	-0.07	0.46	-1.63	0.68	-8.87	0.24	–	–	0.03	0.46
SE 2011	-0.04	0.97	-2.03	0.56	-0.32	0.69	4.71	0.51	0.13	0.25
SPECT 2005	0.10	0.40	1.18	0.64	0.96	0.71	–	–	-0.01	0.16
SPECT 2009	-0.12	0.42	0.39	0.69	-0.76	0.75	–	–	0.01	0.77
CCT 2006	-0.59	0.02	-0.02	1.00	0.49	0.89	17.53	<0.01	0.28	0.28
CCT 2010	-0.06	0.89	-2.42	0.19	–	–	11.69	0.02	0.01	0.45

*Regression Coefficient (Slope). Abbreviations as in tables 1 and 2.

Transthoracic echocardiography (2011 AUC edition).

Ten studies using the 2011 edition of the AUC for TTE were analysed (n=9,211). Because one study, [120] was divided into two, there were 11 studies included in this analysis. The weighted proportion of appropriate tests was 0.85 with substantial heterogeneity among the estimates [95% CI, 0.81 to 0.89, I^2= 96.2%, p<0.0001] (Figure 5- 4).

In this group, a significant positive association was found between appropriateness and year of publication (p=0.01), as well as a strong positive association between inpatient status and appropriateness in the total sample (p<0.0001) (Table 5- 4). There were no improvements in the proportion of appropriate tests among classifiable studies over time (Table 5- 3).

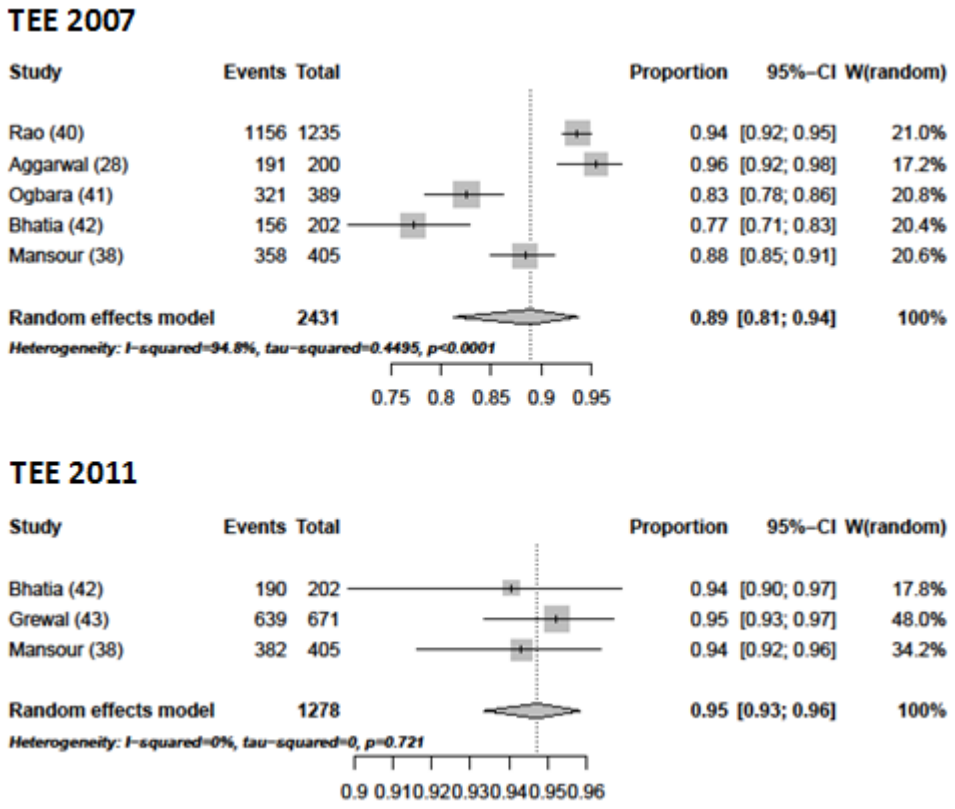
Trans-oesophageal echocardiography.

There were five studies based on the TEE 2007 AUC (n=2,431) published between 2009 and 2012 (Figure 5- 5). The weighted proportion of appropriate tests was 0.89 with substantial heterogeneity among estimates [95% CI, 0.81 to 0.94, I^2= 94.8%, p<0.0001]. In the overall sample, there were negative associations between appropriateness, publication year (p=0.03), and inpatients (p=0.007) (Table 5- 4).

There were no significant associations between appropriateness and gender, ordering physician, or age. There was an apparent correlation between the proportion of classified studies and publication year (Table 5- 3).

Three studies used TEE 2011 AUC, which including 1,278 tests. There was no association with time, gender, specialists or type of patients.

Figure 5- 5 Appropriateness of TEE using the 2007 (up) and 2011 (down) AUC edition

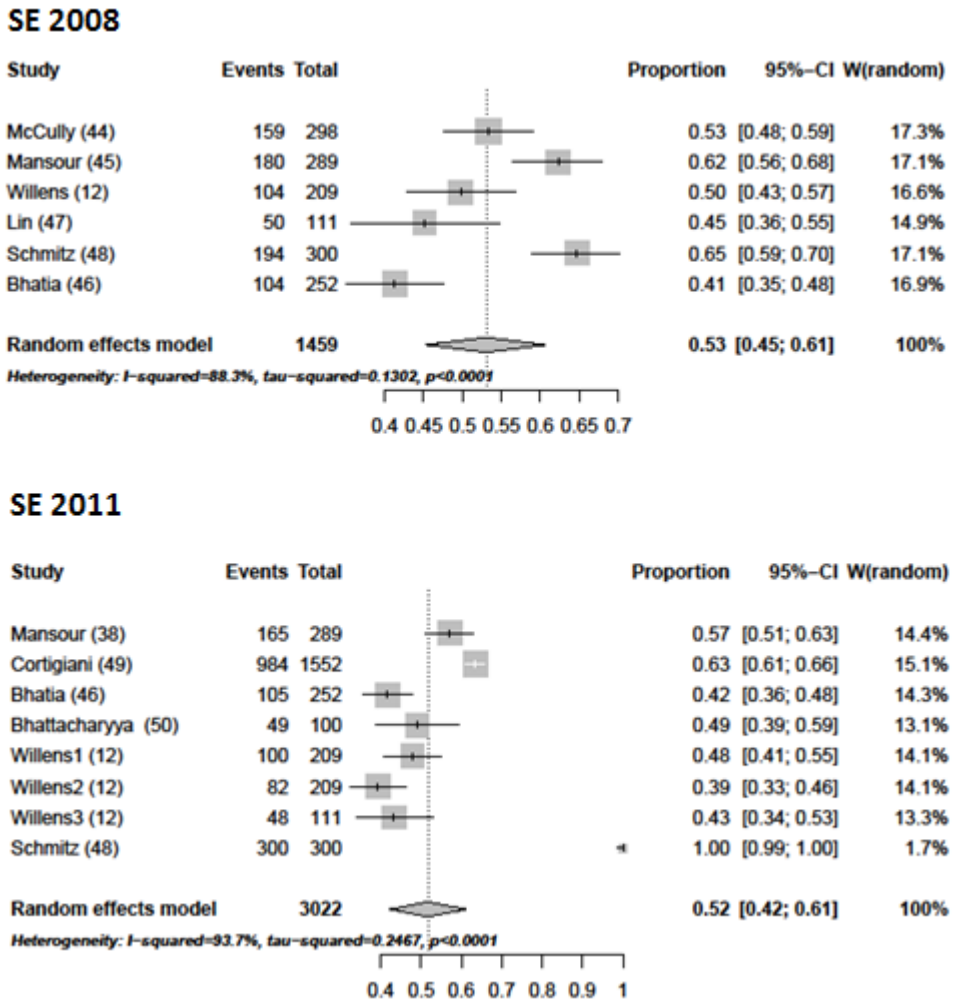


Studies were ordered by increasing publication year. The overall estimate was calculated from random effect model.

Stress echocardiography.

The 2008 AUC were used in six studies (n=1,459) published between 2009 and 2013, with enrollment data between 2005 and 2011. The pooled appropriate testing proportion using the 2008 AUC for SE was 0.53 with significant heterogeneity between estimates [95% CI: 0.45 to 0.61; I^2= 88.3%, p<0.0001] (Figure 5- 6).There were no significant associations between the proportion of appropriateness and publication year, gender, or specialists for either all imaging studies or classifiable studies (Table 5- 3 and Table 5- 4).

Figure 5- 6 Appropriateness of SE using the 2008 (up) and 2011 (down) AUC edition



Studies were ordered by increasing publication year. The overall estimate was calculated from random effect model.

In regards to the analysis of the AUC 2011 for SE, one study[121] was divided into three. Thus, eight studies with a total of 3,022 tests were included (Figure 5- 6). The average appropriateness was 0.52 [95% CI: 0.42 to 0.61; I^2= 93.7% p< 0.0001]. No significant associations between appropriateness and publication year, gender, or specialists were found (Table 5- 4) either among all studies or classifiable studies (Table 5- 3).

SPECT.

Ten studies used the 2005 AUC version for SPECT, with one study[136] divided in two resulting in 11 studies (n=12,694 tests). The weighted proportion of appropriate tests was 0.72 [95% CI:

0.66 to 0.77; I^2= 97.2%, p<0.0001] (Figure 5- 7). No significant associations between appropriateness and publication year, gender, or specialists were found (Table 5- 4).

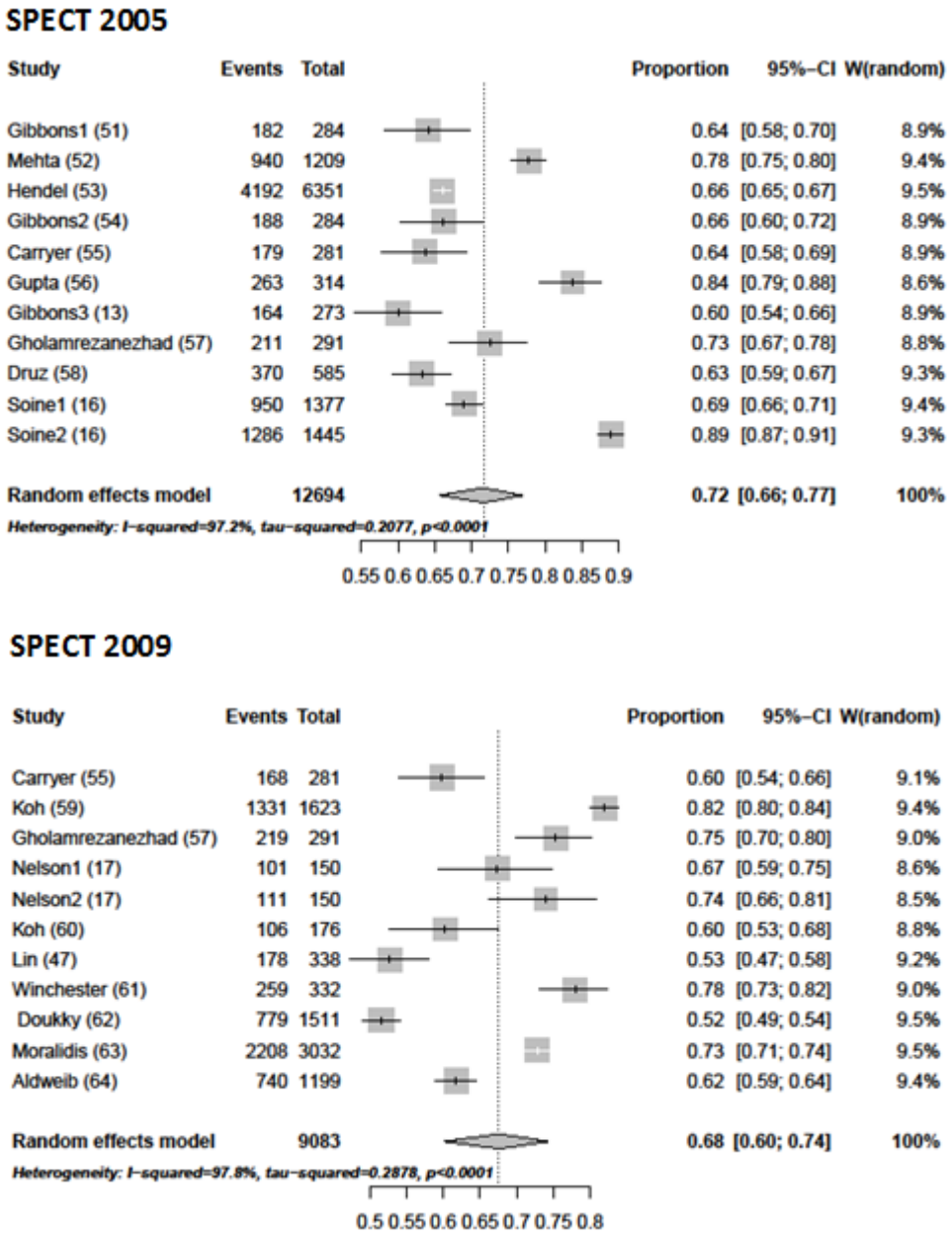
Using the 2009 AUC edition for SPECT, ten studies were found (n=9,083 tests), with one study divided in two. The weighted proportion of appropriate tests was 0.68 [95% CI: 0.60 to 0.74; I^2= 97.8%, p <0.0001]. No significant associations between appropriateness and publication year, gender, or specialists were found.

CTA.

Eight studies (n=29,140) were found evaluating the 2006 AUC for CTA, with one paper [144] divided in two. The average of appropriateness was 0.37 [95% CI: 0.21 to 0.55; I^2 = 99.6%, p<0.0001] (Figure 5- 8). A drop in the proportion of appropriate tests in relation to the year of publication of the paper was found (p=0.02) (Table 5- 4). No significant associations between appropriateness and gender, or specialists were identified. However, a strong positive association between the proportion of appropriateness and hospitalised patients (inpatients) was found (p <0.0001).

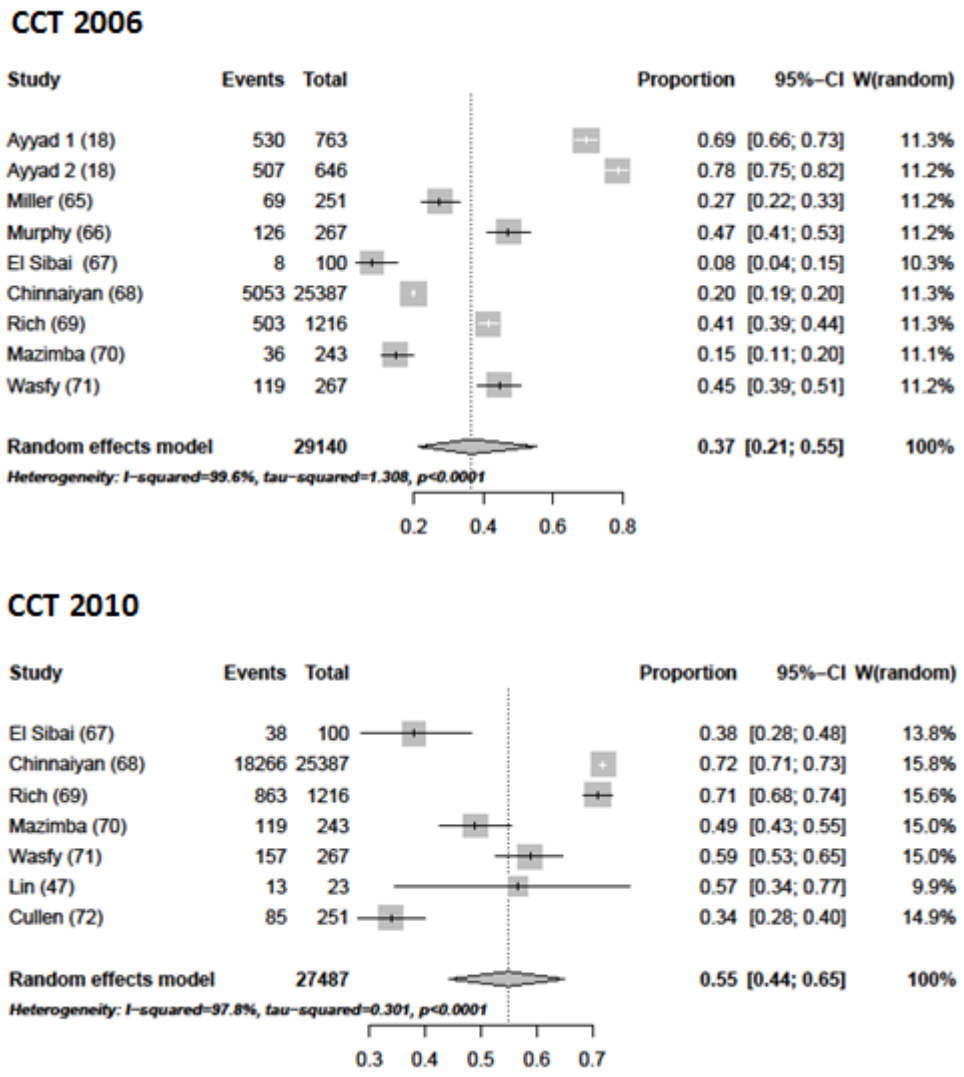
For the evaluation of the 2010 AUC for CTA, seven studies included 27,487 tests (Figure 5- 8). The weighted proportion of appropriate tests was 0.55 [95% CI: 0.44 to 0.65, I^2= 97.8% p<0.0001]. Only a significant association between appropriateness and inpatient tests was found (p=0.0206) (Table 5- 4). There was a significant diminution of classified studies in relation to publication year (Table 5- 3).

Figure 5- 7 Appropriateness of SPECT using the 2005 (up) and 2009 (down) AUC edition



Studies were ordered by increasing publication year. The overall estimate was calculated from random effect model.

Figure 5- 8 Appropriateness of CCT using the 2006 (up) and 2010 (down) AUC edition



Studies were ordered by increasing publication year. The overall estimate was calculated from random effect model.

"Rarely appropriate" (inappropriate) and "maybe appropriate" (uncertain) tests.

Table 5- 5 and Table 5- 6 present the proportion of "rarely appropriate" (inappropriate) and "maybe appropriate" (uncertain) tests, respectively, and their association with the publication year in the total sample and among classifiable imaging studies.

There was a significant diminution of 'rarely appropriate tests' in the 2007 TEE AUC with time. However, no such improvement was noted with the new AUC edition. There was no

impact of the AUC in 'rarely appropriate tests' for other cardiac imaging modalities. Between the old and new editions, "rarely appropriate" stress echocardiography studies increased from 18% to 27%, SPECT from 11% to 20%, and CTA from 17% to 21%.

TEE 2007 showed a decrease in the 'maybe appropriate tests' through time (beta=-0.84, p=0.04). CTA 2006 showed an increase in the proportion of 'maybe appropriate tests' of classified imaging studies but not of the total sample, over time (beta=0.31, p=0.01). No other improvements in the 'maybe appropriate tests' were found for the remaining modalities.

Table 5- 5 Meta-regression of proportion of "Rarely appropriate" tests (Inappropriate) of total sample and classified studies as dependent variables

	Total Sample			**Classified Studies**		
	% "Rarely appropriate"	Pub year*	p value	% "Rarely appropriate"	Pub year*	p value
TTE 2007	0.08 [0.06;0.11]	-0.25	0.06	0.09 [0.07;0.13]	-0.26	0.06
TTE 2011	0.09 [0.06;0.12]	-0.63	0.09	0.09 [0.07;0.12]	-0.64	0.08
TEE 2007	0.01 [0.01;0.03]	-0.45	0.00	0.01 [0.01;0.03]	-0.43	0.00
TEE 2011	0.02 [0.02;0.04]	–	–	0.03 [0.02;0.04]	–	–
SE 2008	0.18 [0.11;0.27]	-0.03	0.91	0.23 [0.16;0.32]	0.02	0.92
SE 2011	0.27 [0.22;0.33]	0.10	0.83	0.28 [0.22;0.35]	0.02	0.97
SPECT 2005	0.11 [0.09;0.14]	-0.13	0.28	0.12 [0.10;0.14]	-0.15	0.21
SPECT 2009	0.20 [0.14;0.28]	0.05	0.81	0.20 [0.14;0.28]	0.05	0.78
CCT 2006	0.17 [0.12;0.23]	0.40	0.09	0.21 [0.13;0.33]	0.45	0.05
CCT 2010	0.21 [0.11;0.37]	-0.55	0.43	0.23 [0.12;0.39]	-0.40	0.56

Values in brackets are 95% confidence intervals. *Regression coefficient (slope). Abbreviations as in Tables 1 and 2.

Table 5- 6 Meta-regression of proportion of "May be appropriate" (uncertain) tests of total sample and classified studies as dependent variables.

	Total Sample			**Classified Studies**		
	% May be appropriate	Publication year*	p value	% May be appropriate	Publication year*	p value
TTE 2007	–	–	–	–	–	–
TTE 2011	0.04 [0.03;0.06]	-0.84	0.05	0.04 [0.03;0.06]	-0.84	0.04
TEE 2007	–	–	–	–	–	–
TEE 2011	0.01 [0.00;0.03]	–	–	0.01 [0.00;0.03]	–	–
SE 2008	0.06 [0.03;0.09]	-0.28	0.26	0.08 [0.05;0.12]	-0.22	0.31
SE 2011	0.15 [0.09;0.24]	0.02	0.98	0.16 [0.10;0.24]	0.07	0.95
SPECT 2005	0.11 [0.09;0.14]	0.09	0.35	0.12 [0.10;0.15]	0.07	0.50
SPECT 2009	0.08 [0.05;0.13]	0.06	0.81	0.09 [0.05;0.14]	0.07	0.80
CCT 2006	0.19 [0.12;0.29]	0.26	0.17	0.24 [0.18;0.31]	0.31	0.01
CCT 2010	0.11 [0.08;0.16]	0.12	0.78	0.12 [0.09;0.17]	0.24	0.59

Values in brackets are 95% confidence intervals. *Regression coefficient (slope). Abbreviations as in Tables 1 and 2

Discussion

This comprehensive assessment of the published literature evaluating AUC for different cardiac imaging techniques from 103,567 tests grouped in 10 different cardiac imaging outcomes. Meta-regression was used to assess the temporal trend of appropriateness, based on the year of publication. In contrast to studies comparing the behaviour of specific groups of physicians over time, the results are an indication of "real-world" practice at sites publishing their appropriate use data.

There are five significant findings; first, the improvement of appropriate use from the original to the revised versions may merely reflect an easier classification of patients and a change in attribution of the proportion of appropriate testing rather than a change in practice. Second, this study showed a temporal improvement in percent appropriateness for TTE, TEE and CTA, but no evidence of a change in the number and proportion of appropriate testing for other modalities. Overall rates of appropriate use for CTA and SE remain low, and those for SPECT only modest. This implies a disconnection between clinical practice and AUC that warrants better understanding. Third, this limited change has not matched the reduction of imaging tests over the last five years, suggesting that physician use of AUC in the ordering process may not have played a significant role in this decrease. However, an indirect role (though AUC influence on the decisions of Radiology Benefit Managers (RBM)) cannot be excluded. Fourth, the proportion of appropriate use presented here may well be shown in its best light in these retrospective and largely unblinded evaluations, which for the most part were performed by physicians able to identify appropriate indications, even if this was not the primary reason for the test. Moreover, the studies were mainly made at the point of service of academic medical institutions. There was substantial variation between observers, in particular between physicians. Finally, although there was some evidence of publication bias for SE 2011 and CTA 2010, bias was not identified for the majority of scenarios.

Understanding temporal variations in appropriate use.

The observed heterogeneity of the proportion of appropriate testing among studies might be expected on the basis of a wide variety of participant characteristics, study designs, types of hospital and regions in the published data. However, the unique aspect of this study is its examination of the temporal variation in the proportion of appropriate testing. This is difficult to measure in a specific study as ordering physicians may not behave as they do in daily life.

The drivers of test ordering are complex, and the persistent rate of approximately 80% for TTE (and less for other modalities) perhaps testifies to a variety of influences that are not disease-specific and include factors individual to the patient, including those related to other comorbidities and situational considerations.

These features may drive the request for testing in a situation when the test is considered "rarely appropriate". In addition, clinical practice guidelines may be discordant with AUC. Finally, the adjudication of appropriate testing is often inconsistent because appropriate and rarely appropriate reasons for testing may co-exist in the same patient. The implication is that the ordering physician may choose an existing appropriate indication rather than the real clinical issue. This change in indication is especially likely to happen when the proportion of inappropriate (or rarely appropriate) tests is reviewed as part of the accreditation process [66].

In this respect, the increment of "rarely" appropriate tests from the first to the second versions of stress echocardiography, SPECT, and CTA, was a surprising finding of this meta-analysis. Interestingly, none of these tests showed a gradation of "rarely appropriate" use within the time frame of each edition; therefore, this finding likely reflects the change in criteria rather than a change in practice.

Alternative approaches to reducing cost.

The use of AUC as a process to reduce costs by trying to maintain appropriate testing neglects the fact that testing labeled as "sometimes" and "rarely" appropriate is very appropriate in some situations. Indeed, this is a shortcoming of the widespread use of Radiology Benefit Managers (RBM) as a tool to control the use of cardiac imaging: they are inflexible to situational demands[121]. The application of AUC at point-of-care, for example using electronic tools that help physicians to choose "appropriately" has produced similar results to the RBM, but has the same limitation. [45]

Although the nuances of particular clinical scenarios make the AUC problematic for controlling testing, they are potentially valuable as a yardstick for education. The evidence regarding the value of educational campaigns is currently contradictory [120-122]. In the interpretation of responses to AUC campaigns, it should be kept in mind that knowledge of AUC is but one component of test selection, which is also influenced by the characteristics of health professionals, features of practice settings, incentives, linkage of AUC performance with accreditation or licensing bodies, patient factors, compatibility with existing practice and

beliefs, and perceived quality of the guidelines.[153] Moreover, repetition requirements for an educational campaign to have a sustained effect are unclear.

In jurisdictions where the laboratory is responsible for appropriate use, a strategy of laboratory-based audit is needed for the thousands of cardiac imaging requests which are submitted to the laboratory every year. The use of AUC to facilitate audit is more likely to be effective than its application to individual test requests. Tests that are most likely to be of "maybe" or "rare" appropriateness include those requested in younger patients, those with previous tests, those who are outpatients or tests that are re-evaluations in asymptomatic patients or without changes in clinical status.[59] These situations might be used as markers of potential inappropriate use in individual patients.

Study limitations.

This study has sought to link reported percent appropriateness rates with the reduction of imaging over time, but there are some potential problems. First, most cardiovascular imaging is performed in the community practice environment, while most AUC studies have been reported in academic medical centres.

Nonetheless, the broad discussion of AUC over the last decade might be expected to influence all environments, and a practice-specific variation of some tests and not others seems unlikely. Second, it is unclear whether changes in percent appropriateness reflect better test selection rather than observer-expectancy effect, or coding of indications to satisfy AUC.

Conclusions

Improvements in the percent appropriateness rate of TTE seem to correlate with the temporal reduction in imaging. However, methodological problems in this literature – including possible publication bias and retrospective assignment of AUC ratings in most studies by individual(s) affiliated with the institution where the study was performed – may compromise confidence in this observation. Moreover, these changes are not uniform, with no association between the rate of appropriateness and date of publication for SPECT, TEE, CT and SE. It is possible that the reduction of imaging tests is unrelated to AUC.

Chapter 6. Impact of Appropriate Use Criteria on Survival

The research contained within this chapter has been published as[7]:

- Fonseca R, Otahal P, Galligan J, Neilson S, Huynh Q, Saito M, Negishi K and Marwick TH. Association of Survival Time with Transthoracic Echocardiography in Stable Patients with Heart Failure: Is Routine Follow-Up Ever Appropriate? *International Journal of Cardiology*. 2016.

Preface

As it was discussed in the introductory chapter, the achievement of appropriate use of cardiac imaging has become a crucial battle in an attempt to control imaging use and health expenditure growth. The AUC were designed to achieve these aims. Australia is no stranger to this problem; it was demonstrated that growth in imaging use has affected Australia over the last decade but most importantly, it seems the imaging use, especially echocardiography, is not properly controlled [4].

The interest in using the AUC has become a current topic of discussion in Australia. However, as it has been shown in chapters 3 and 4, there are inconsistencies between the guidelines and AUC which may challenge our physicians and moreover, it seems that, despite all the focus on the use of AUC, these have not achieved the goal of changing the ordering behaviour in doctors, demonstrated by a lack of significant increase in the proportion of "appropriate" testing and the consequential decrease in "inappropriate" screening[5, 6].

Once more, it has been evidenced that the use of transthoracic echocardiography is one of the most challenging in the medical practice: there has been an increase in the proportion of "appropriate" tests over time due to a decrease in the proportion of "uncertain" tests. However, the proportion of "inappropriate" tests has not changed[6].

In this chapter, we sought to determine if the AUC have an impact on health outcomes. In the AUC for echocardiography, most of the inappropriate scenarios are defined by timing of follow-up testing and absence or change of symptoms or signs of cardiac disease (asymptomatic or stable patients)[26] and there are some indications that are uncertain due to a lack of scientific evidence.

For this reason, we have chosen to analyse the impact of the criteria on heart failure survival and readmission and to provide scientific evidence to the AUC. Heart failure is one of the most expensive syndromes for any health system due to the poor quality of life, high prevalence and incidence, and costly care[154, 155]. In Australia, the cost of chronic heart failure has been estimated at more than $ 1 billion per year[156], and the increase in readmissions and mortality have increased during the last decade[157]. This burden on the health care system can be improved by reducing the time to and rates of readmission and death through the appropriate use of cardiac imaging.

Abstract

Background. The appropriateness of repeat transthoracic echocardiography (TTE) for stable heart failure (HF) is based on timing of the follow-up examination, but this lacks scientific support. We sought the association of routine follow-up TTE on survival and readmission in stable HF.

Methods. Patients with HF were selected from consecutive HF admissions from 2008-2012. Groups were divided into: no follow-up TTE; routine <1 year with no change in status ("inappropriate"), ≥1year follow-up with no change in status ("uncertain") and TTE due to change in clinical status ("appropriate"). Survival analysis was performed for the combined endpoint of HF readmission and death, and a separate analysis was performed for HF readmission, with death as a competing risk.

Results. Of 550 HF patients, 141 had a follow-up TTE, including 41 (29%) within 1 year. The event-free time was 1.10 years [95% CI: 0.69, 1.49] for no TTE, 2.61 years [95% CI: 1.08, 3.04] for the "inappropriate" group, 2.45 years [95% CI: 1.37, 5.78] for the "uncertain" group, and 0.09 years [95% CI: 0.02, 1.80] for the "appropriate" group ($p<0.001$ between all groups; $p=0.16$ between "inappropriate", "uncertain" and "appropriate" test groups; $p=0.06$ between "inappropriate" and "uncertain" groups). HF readmission was not associated with follow-up TTE timing. There were no differences in the cumulative incidence for death between groups. There were no differences in change in management in "inappropriate" and "uncertain" tests.

Conclusion: The distinction of appropriateness of routine repeat TTE in stable HF patients, based on testing <1 or ≥1 year after index admission appears unjustified.

Keywords: appropriate use, heart failure, echocardiography

Introduction

The Appropriate Use Criteria (AUC) for transthoracic echocardiography (TTE) were designed to facilitate selection of the most appropriate testing for individual clinical situations, in the context of rational use, standardization of clinical practice, and delivery of high-quality care [78]. However, there have been concerns that the determination of AUC on the basis of expert opinion carries the risk of non-scientific guidance [5, 64, 103]. This is particularly problematic when the only distinction between an "appropriate" and an "inappropriate" (also called "rarely appropriate") TTE is the time of routine follow-up evaluation or the symptom status of the patient [78].

Recent studies have questioned the relationship between appropriateness assessment and the clinical impact of testing [35, 158]. Although there is no a perfect way to determine clinical impact of a TTE, its benefits have been measured by the change in management and surprisingly, an average of one third of appropriate TTE have led to a change in management, with no differences between "appropriate" or "inappropriate" TTE [35].

Heart failure (HF) is one of the most expensive and highly prevalent cardiovascular conditions. The evidence regarding follow-up and clinical benefit of the routine use of TTE in stable patients with HF is sparse [78]. Additionally, there have not been studies of the association between survival in HF and appropriate use status. Therefore, we sought to assess the association between survival and time to HF readmission and appropriate use in HF patients with no change in clinical status or cardiac exam.

The main objectives of this study were to assess:

1) Association of routine follow-up tests with survival and HF readmission in stable HF,

2) Differences in outcome between tests performed with <1 year follow-up (inappropriate, or "rarely appropriate"), with ≥1 year follow-up (uncertain appropriateness), "appropriate" TTE (those performed in response to a change in status or exam) or no follow-up TTE,

3) Associations of changes in management after TTE on survival and HF readmission time in each described AUC group,

4) Role of routine follow-up TTE in stable HF patients, more specifically "uncertain" tests compared to "inappropriate" tests.

Methods

Study design and inclusion criteria.

This cohort analysis analysed event-free survival, cause-specific hazard, and cumulative incidence involving the subsequent outcomes (HF readmission or death) of HF patients. The study included available data of all patients hospitalised for an initial heart failure admission at a tertiary referral hospital between 1st July 2008 and 30 June 2012.

Patients were included if they were ≥18 years old and had a diagnosis of HF at discharge and had a TTE before or during their first HF admission. Patients were excluded if they were <18 years old or did not have a TTE previously or during the HF admission.

A routine follow-up TTE was defined as a TTE performed for a periodic evaluation in a patient with no change in clinical status or cardiac exam within a year (inappropriate) or after a year (uncertain appropriateness). "Appropriate" TTE in these HF patients were defined as being due to a change in clinical status or cardiac exam. "Change in clinical status" was defined as any change in the stability of the condition determined by a new onset or worsening of signs or symptoms of left-sided failure, right-sided failure or biventricular failure or apparition of new symptoms or signs such as chest pain, shock, hypotension, syncope, arrhythmia, murmur, or peripheral embolic event.

"Change in cardiac exam" corresponded to any change in observation, palpation or auscultation during cardiac exam. "Stable HF" was defined as a patient with HF with no visible progression of the disease, determined by no change in symptoms or signs of left or right-sided failure or biventricular failure.

Change in management.

Change in management was determined as any change of care occurred in response to the follow-up TTE. These changes included changes in medications (beta-blockers, angiotensin-converting-enzyme inhibitor/angiotensin-II receptor blockers, diuretics, inotropes, anticoagulants, and mineralocorticoid antagonists), referral to another subspecialty, surgery or invasive procedures, cardioversion, or new diagnostic testing or change/cancellation of the initially planned management. A complete review of the electronic medical record (EMR) was performed for each TTE. Due to the possible permutations, "change in management" was not studied in detail and it was analysed as a binary variable (yes/no).

Clinical characteristics.

General characteristics and baseline comorbidities were identified from the EMR. These variables (Appendix Table 5) included demographics, systolic blood pressure, functional status (New York Heart Association Functional Classification, NYHA), medications, and laboratory findings (creatinine (µmol/L), left ventricular ejection fraction (LVEF, %), medical history and comorbid disease. Risk of death was determined for each patient using the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) HF risk calculator [159].

Follow-up TTE.

Four groups were defined, in accordance with routine follow-up time described in the AUC:

- "Appropriate": patients with follow-up TTE due to a change in clinical status or cardiac exam,
- "Inappropriate": routine follow-up <1 year (defined as rarely-appropriate by the AUC),
- "Uncertain": routine follow-up ≥1 year [78] ,
- No-TTE: patients with no follow-up TTE.

The starting time was defined as the day of discharge.

Outcomes and end points.

The primary outcomes were the combined endpoint of HF readmission or death from any cause and HF readmission and death separately. HF readmission was defined as any subsequent hospital admission for which the primary diagnosis was recorded as HF; time to HF readmission was determined from the date of discharge of first hospital admission for HF to the day of the readmission for HF. Time to death was recorded as the time between the date of hospital discharge until the date of death. Participants who were free of events (HF readmission or death) were censored when follow-up time ended.

Statistical analysis.

This survival analysis comprised event-free survival, cause-specific hazard, and cumulative incidence. All statistical analyses were performed using R software [35]. Continuous variables are reported as mean (SD). Categorical variables are expressed as proportions. Non-normally distributed continuous variables are expressed as median and interquartile range (IQR).

Baseline differences between all groups were detected using ANOVA for normally distributed continuous variables and the nonparametric Kruskal-Wallis rank sum test for non-normally distributed continuous variables. A t-test was used to compare variables between the two groups and the Mann-Whitney U test was used for non-normally distributed variables. The Fisher's exact test was used for categorical variables with less than 5 observations, and a chi-squared test for other similar variables.

For the event-free survival, the analysis was focused on event-free status, plotted with the Kaplan-Meier curves. Cox proportional hazard regression was used to assess the event-free survival for a combined endpoint (HF readmission or death). Models were assessed by testing the proportional hazards assumption and by examining Schoenfeld residuals plots.

For the competing risk analysis, HF readmission was defined as the main outcome and death as the competing event. In both analyses follow-up TTE was treated as a time-varying covariate; with patients contributing to the risk-set in the No-TTE group until the date of their follow-up, upon which these patients switched to contribute to the risk-set in either the "appropriate", "inappropriate" or "uncertain" groups (whichever was applicable). Multivariable analysis were performed using a purposeful selection of covariates [160].

The study was performed on available administrative data of all patients hospitalised for initial heart failure admission at a tertiary hospital between 1st July 2008 and 30 June 2012. On the basis of 75% of the patients having no follow-up TTE, and an anticipated 1 year survival free of readmission and death in this group, the study had an 80% power at an alpha of 0.05 to identify a difference in outcome with non-TTE and TTE groups of 302 and 101, respectively [161].

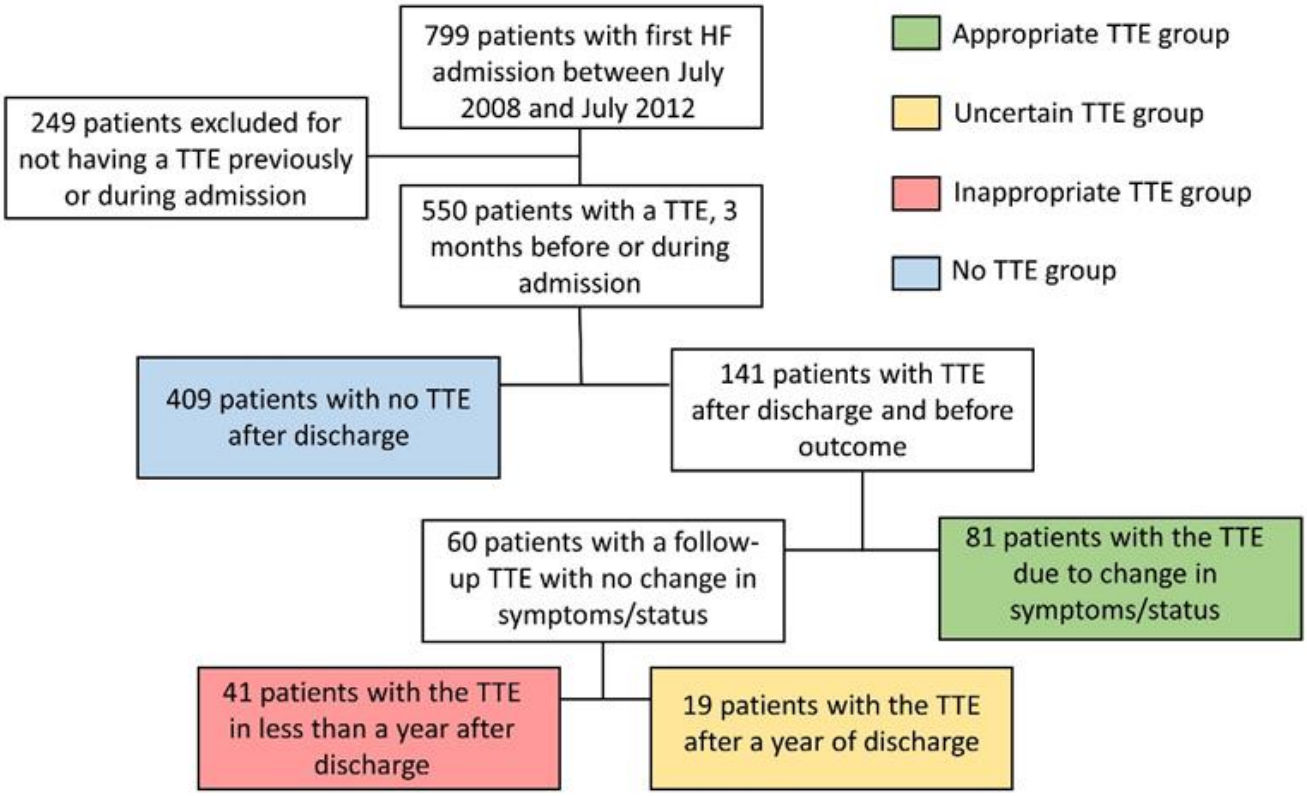
Results

Study population.

The study population comprised 550 (69%) of 799 patients with a first Heart Failure admission, identified from July 2008 to June 2012, who had TTE prior to or during admission. Of the 141 in this group who had a TTE after discharge and before HF readmission, 60 had a routine ("inappropriate" or "uncertain") follow-up TTE, performed for a periodic evaluation with no change in clinical status or cardiac exam, and 81 patients had an "appropriate" second TTE due to change in clinical status or cardiac exam.

The 60 patients who had a routine follow-up TTE were divided in two groups according to AUC: 41 patients undertaking routine follow-up in <1 year ("inappropriate") and 19 with ≥1-year routine follow-up ("uncertain" by the AUC). The remaining 409 patients had no follow-up TTE (Figure 6- 1).

Figure 6- 1 Flow chart of the study



The general characteristics of population are shown in Appendix Table 5. In general, patients with no follow-up TTE were older than the other groups, and were also more likely to have higher levels of creatinine, greater risk of death at 1 and 3 years (measured by the MAGGIC score) and more likely to have NYHA class 3 and 4 HF. There were no differences between people with "inappropriate" and "uncertain" tests, except for a higher proportion of history of myocardial infarction in the "uncertain" TTE group (58% vs 27%, p=0.04) (Table 6- 1).

Table 6- 1 Baseline characteristics between no follow-up TTE and follow-up TTE groups (“inappropriate”, “uncertain” and “appropriate” groups).

	No TTE	**TTE**	**p**
n	409	141	
age (median [IQR])	79 [70, 84]	74 [66, 79]	<0.01
Gender male, n (%)	222 (54.3)	85 (60.3)	0.25
Body mass index, median [IQR]	26 [23, 30]	27 [24, 32]	0.01
Systolic blood pressure in mmHG, median [IQR]	120 [105, 135]	120 [110, 136]	0.31
Creatinine levels in µ/L, median [IQR]	112 [86, 148]	96 [74, 120]	<0.01
Left ventricular ejection fraction %, median [IQR]	0.45 [0.32, 0.60]	0.45 [0.35, 0.58]	0.69
MAGGIC score, median [IQR])	28 [23, 33]	24 [20, 28]	<0.01
Risk of death 1 year, median [IQR]	0.21 [0.13, 0.32]	0.15 [0.10, 0.21]	<0.01
Risk of death 3 years, median [IQR]	0.46 [0.32, 0.63]	0.34 [0.25, 0.46]	<0.01
Betablockers, n (%)	225 (55.0)	90 (63.8)	0.08
Diuretics, n (%)	369 (90.2)	122 (86.5)	0.29
Mineralocorticoids, n (%)	124 (30.3)	49 (34.8)	0.38
Angiotensin-converting-enzyme inhibitor/ angiotensin-II receptor blockers, n (%)	304 (74.3)	106 (75.2)	0.93
Calcium antagonists, n (%)	82 (20.0)	24 (17.0)	0.51
Antiarrhythmics, n (%)	44 (10.8)	17 (12.1)	0.79
Digoxin, n (%)	84 (20.5)	27 (19.1)	0.82
Statins, n (%)	210 (51.3)	74 (52.5)	0.89
Hypertension, n (%)	292 (71.4)	101 (71.6)	1.00
Dyslipidaemia, n (%)	194 (47.4)	77 (54.6)	0.17
History of angina, n (%)	119 (29.1)	55 (39.0)	0.04
History of atrial fibrillation, n (%)	203 (49.6)	75 (53.2)	0.53
Arrhythmia, n (%)	68 (16.6)	26 (18.4)	0.72
Cardiomyopathy, n (%)	153 (37.4)	48 (34.0)	0.54

Dilated cardiomyopathy, n (%)	115 (28.1)	38 (27.0)	0.88
Hypertrophic cardiomyopathy, n (%)	25 (6.1)	8 (5.7)	1.00
Restrictive cardiomyopathy, n (%)	13 (3.2)	2 (1.4)	0.38
Deep vein thrombosis, n (%)	20 (4.9)	5 (3.5)	0.67
Angioplasty, n (%)	78 (19.1)	33 (23.4)	0.33
Myocardial infarction, n (%)	178 (43.5)	52 (36.9)	0.20
Cerebrovascular disease, n (%)	93 (22.7)	31 (22.0)	0.95
Renal disease, n (%)	106 (25.9)	31 (22.0)	0.41
Valvular disease, n (%)	152 (37.2)	51 (36.2)	0.91
Cardiac catheterisation, n (%)	53 (13.0)	19 (13.5)	0.99
Diabetes mellitus, n (%)	166 (40.6)	53 (37.6)	0.60
Chronic obstructive pulmonary disease, n (%)	167 (40.8)	61 (43.3)	0.69
NYHA class, n (%)			<0.01
1	78 (19.1)	36 (25.5)	
2	111 (27.1)	49 (34.8)	
3	142 (34.7)	48 (34.0)	
4	78 (19.1)	8 (5.7)	
NHYA class 3 and 4, n (%)	220 (53.8)	56 (39.7)	0.01
Active change in management after follow-up TTE (%)	_	30 (21.3)	_

Abbreviations: TTE: transthoracic echocardiography; p: p value (significant values <0.05); IQR: Inter quartile range; mmHg: millimeters of mercury; µ/L: micromoles per litre; NYHA: New York Heart Association heart failure classification.

In general, patients with "inappropriate" tests were younger, had more favourable levels of creatinine, had worse ejection fraction and less history of hypertension, dyslipidaemia, and angina. No differences were found in clinical status, risk of death and HF readmission or mortality.

The "uncertain" test group was similar to the "appropriate" test group except in history of cardiovascular disease, which was higher in the "appropriate" test group. No differences were found in proportion of outcomes between these groups.

Event free survival analysis.

The median follow-up time was 1.02 years (range 0.003, 6.48) and the median time for follow-up TTE was 0.57 years (range 0.01, 5.00). There was a statistically significant difference between the proportions of patients with HF readmission between the groups (Table 6- 2). There were no significant differences in mortality between groups.

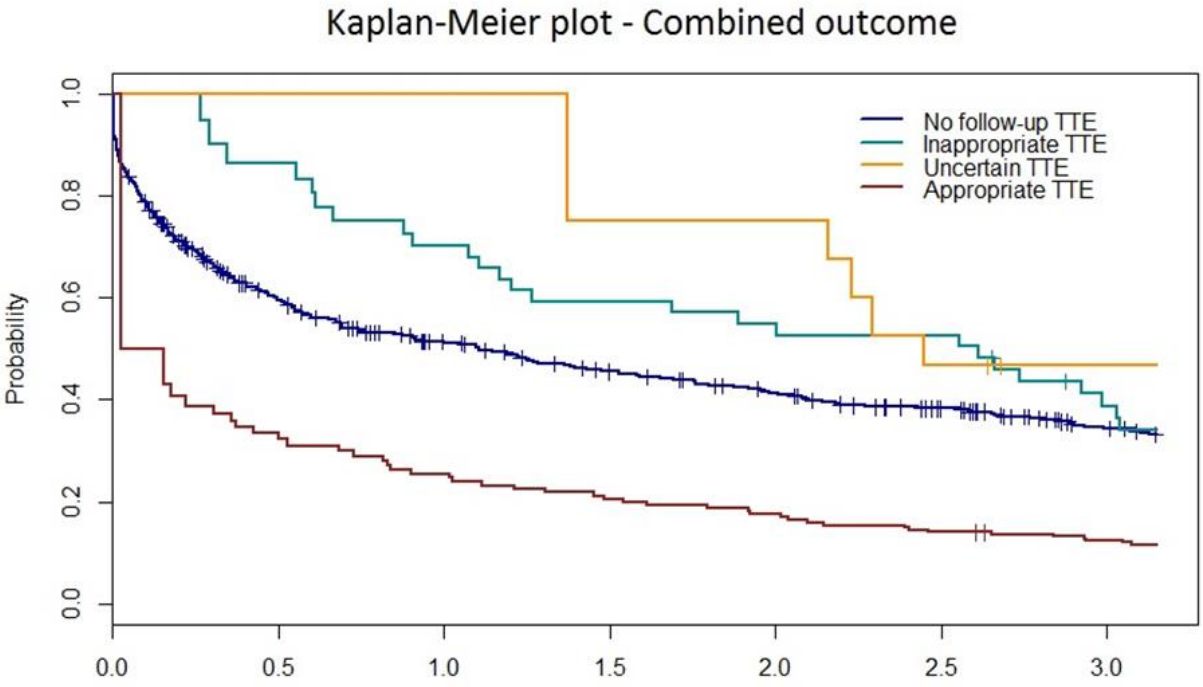
Table 6- 2 Outcomes in no follow-up TTE and follow-up TTE groups (Inappropriate, "uncertain" and "appropriate" groups).

	HF readmission, n (%)	**Death, n (%)**	**HF/death, n (%)**
TTE	28 (19.9)	63 (44.7)	91 (64.5)
No TTE	166 (40.6)	170 (41.6)	336 (82.2)
p (TTE vs none)	<0.01	0.58	<0.01
Inappropriate	10 (24.4)	17 (41.5)	27 (65.9)
Uncertain	1 (5.3)	8 (42.1)	9 (47.4)
Appropriate	17 (21.0)	38 (46.9)	55 (67.9)
p (all groups)	<0.01	0.85	<0.01
p (inappropriate, uncertain and appropriate groups)	0.21	0.82	0.24
p (inappropriate and uncertain groups)	0.15	1	0.28
p (inappropriate and appropriate groups)	0.84	0.71	0.98
p (uncertain and appropriate groups)	0.18	0.9	0.16

Abbreviations: HF=heart failure, TTE=transthoracic echocardiography.

Figure 6- 2 shows the survival function for HF readmission and death as combined outcome in each of the TTE groups. The estimated median survival time was 1.10 years [95%CI: 0.69, 1.49] for the no TTE group, 2.61 years [95% CI: 1.08, 3.04] for the "inappropriate" test group, 2.45 years [95% CI: 1.37, 5.78] for the "uncertain" test group, and 0.09 years [95% CI: 0.02, 1.80] for the "appropriate" test group (log rank test 23.1, p<0.001 between all groups; log rank test 3.7, p=0.16 between "inappropriate", "uncertain" and "appropriate" test groups, and log rank test 3.5, p=0.06 for "inappropriate" and "uncertain" groups).

Figure 6- 2 Survival function for HF readmission and mortality as combined outcome in patients with HF after discharge from first HF admission.



Numbers at risk

	0	30 days	60 days	0.5	1	1.5	2	2.5	3
No TTE	0	426	377	274	215	178	155	131	96
Inappropriate	0	9	13	28	32	27	25	24	17
Uncertain	0	0	0	0	0	4	9	9	8
Appropriate	0	6	16	28	31	34	30	31	32

Estimations of the average effect of risk factors on the overall survival (HF readmission and death as composite outcome) are shown inTable 6- 3. In the univariable hazards models, the event rate among older people, people with higher MAGGIC score (therefore higher risk of death), myocardial infarction, renal disease and NYHA class 3 and 4 was higher than people without these features. People with higher body mass index had a lower event rate than people with lower BMI (Table 6- 3).

Although there was a possible increase in the risk of having HF readmission or death in the "appropriate" TTE group (HR 1.50 [95% CI: 0.95, 2.39], p=0.08) compared to the "inappropriate" group and a similar trend in the "uncertain" group (HR 1.36 [95% CI: 0.63, 2.94], p=0.43) vs the "inappropriate" group, there were no significant differences between groups compared to "inappropriate" tests group as the reference.

The results for the adjusted model show HF readmission to be associated with renal disease and higher HF classification (Table 6- 3). Interestingly there were no differences between TTE groups. Routine follow-up TTE time (<1 year and ≥1 year, or "inappropriate" and "uncertain" respectively), did not show differences in outcome between both groups.

Table 6- 3 Estimations of the average effect of risk factors on the overall survival (combined endpoint)

	Univariable Cox Model		**Multivariable Cox Model**	
	HR (95% CI)	p	HR (95% CI)	p
Inappropriate	Ref	_	Ref	_
Uncertain	1.36 (0.63 - 2.94)	0.43	1.33 (0.61 - 2.86)	0.47
Appropriate	1.50 (0.95 - 2.39)	0.08	1.56 (0.98 - 2.48)	0.06
No TTE	0.93 (0.62 - 1.39)	0.72	0.92 (0.62 - 1.38)	0.70
Age	1.02 (1.01 - 1.02)	<0.01		
Gender male	1.16 (0.96 - 1.40)	0.14		
Body mass index	0.98 (0.96 - 1.00)	0.01		
Systolic blood pressure in mmHG	1.00 (0.99 - 1.00)	0.10		
Creatinine levels in µ/L	1.00 (1.00 - 1.00)	0.07		
Left ventricular ejection fraction %	0.87 (0.44 - 1.70)	0.68		
MAGGIC score	1.06 (1.04 - 1.07)	<0.01		
Risk of death 1 year	20.65 (9.66 - 44.16)	<0.01		
Risk of death 3 years	6.98 (4.19 - 11.62)	<0.01		
Betablockers use	0.84 (0.70 - 1.02)	0.08		
Diuretics use	0.96 (0.71 - 1.30)	0.81		
Mineralocorticoids	1.09 (0.89 - 1.33)	0.42		
Angiotensin-converting-enzyme inhibitor/ angiotensin-II receptor blockers	0.82 (0.67 - 1.02)	0.08		
Calcium antagonists	0.94 (0.74 - 1.19)	0.59		
Antiarrhythmics	1.09 (0.81 - 1.47)	0.57		
Digoxin	1.01 (0.80 - 1.28)	0.91		
Statins	1.03 (0.85 - 1.24)	0.77		
Hypertension	1.05 (0.85 - 1.29)	0.66		

Dyslipidaemia	1.09 (0.90 - 1.31)	0.40		
History of angina	0.96 (0.78 - 1.18)	0.68		
History of atrial fibrillation	1.09 (0.90 - 1.32)	0.37		
Arrhythmia	1.04 (0.81 - 1.32)	0.78		
Cardiomyopathy	0.99 (0.81 - 1.21)	0.92		
Dilated cardiomyopathy	0.96 (0.78 - 1.19)	0.72		
Hypertrophic cardiomyopathy	1.13 (0.77 - 1.65)	0.54		
Restrictive cardiomyopathy	0.94 (0.52 - 1.71)	0.84		
Deep vein thrombosis	1.19 (0.75 - 1.88)	0.47		
Angioplasty	0.87 (0.68 - 1.11)	0.27		
Myocardial infarction	1.27 (1.05 - 1.54)	0.02		
Cerebrovascular disease	1.10 (0.88 - 1.38)	0.38		
Renal disease	1.45 (1.17 - 1.80)	<0.01	1.47 (1.18 - 1.82)	<0.01
Valvular disease	1.19 (0.98 - 1.45)	0.08		
Cardiac catheterisation	0.98 (0.73 - 1.31)	0.88		
Diabetes mellitus	0.90 (0.74 - 1.09)	0.28		
Chronic obstructive pulmonary disease	1.11 (0.92 - 1.35)	0.27		
NHYA class 3 and 4	1.62 (1.33 - 1.96)	<0.01	1.69 (1.40 - 2.06)	<0.01
Active change in management after follow-up TTE	0.79 (0.54 - 1.17)	0.24		

Analysis of heart failure readmission in the presence of mortality as the competing event.

The univariable and multivariable subdistribution hazard ratios of risk factors are shown inTable 6- 4, with timing of routine follow-up TTE shown as a time-dependent variable. The univariable model showed no statistical differences in the risk of HF readmission between the different TTE groups in the presence of death as a competing risk. The MAGGIC score, history of renal disease, coronary disease (as evidenced by use of statins, and cardiac catheterization), lower ejection fraction and NYHA class 3 and 4, were all associated with HF readmission.

The use of statins, history of renal disease, cardiac catheterization and higher NYHA class were the independent predictors of HF readmission (Table 6- 4). Interestingly, there were no differences between TTE groups. No statistically significant differences were found between TTE groups for death as the endpoint in the presence as HF readmission as the competing event (Table 6- 4).

Table 6- 4 Competing risk models for HF readmission in the presence of death as the competing event (univariable/multivariable HF), and for death in the presence of HF readmission as the competing risk (univariable/multivariable mortality).

	Heart Failure				**Mortality**			
	Univariate		Multivariate		Univariate		Multivariate	
	HR (95% CI)	p	HR (95% CI)	p	HR (95% CI)	p	HR (95% CI)	p
Inappropriate	Ref	_	Ref	_	Ref	_	Ref	_
Uncertain	0.20 (0.03 - 1.46)	0.11	0.17 (0.02 - 1.25)	0.08	0.92 (0.44 - 1.91)	0.83	0.84 (0.43 - 1.67)	0.62
Appropriate	0.85 (0.40 - 1.77)	0.66	0.71 (0.33 - 1.51)	0.37	1.13 (0.66 - 1.92)	0.66	1.20 (0.70 - 2.09)	0.51
No TTE	1.65 (0.90 – 3.00)	0.10	1.43 (0.77 - 2.63)	0.26	1.01 (0.63 - 1.60)	0.97	0.87 (0.53 - 1.42)	0.58
Age	1.00 (0.99 - 1.01)	0.92			1.03 (1.01 - 1.04)	<0.01	1.02 (1.01 - 1.04)	<0.01
Gender male	1.20 (0.90 - 1.59)	0.22			1.05 (0.81 - 1.36)	0.71		
Body mass index	1.00 (0.98 - 1.02)	0.85			0.97 (0.94 - 0.99)	<0.01		
Systolic blood pressure in mmHG	0.99 (0.99 – 1.00)	0.08			1.00 (0.99 - 1.01)	0.94		
Creatinine levels in µ/L	_	_			1.00 (1.00 – 1.00)	0.15		
Left ventricular ejection fraction %	0.35 (0.13 - 0.96)	0.04			2.58 (1.04 - 6.41)	0.04		
MAGGIC score	1.02 (1.00 - 1.04)	0.03			1.06 (1.04 - 1.08)	<0.01		
Risk of death 1 year	3.54 (1.16 - 10.8)	0.03			18.60 (6.47 - 53.60)	<0.01		

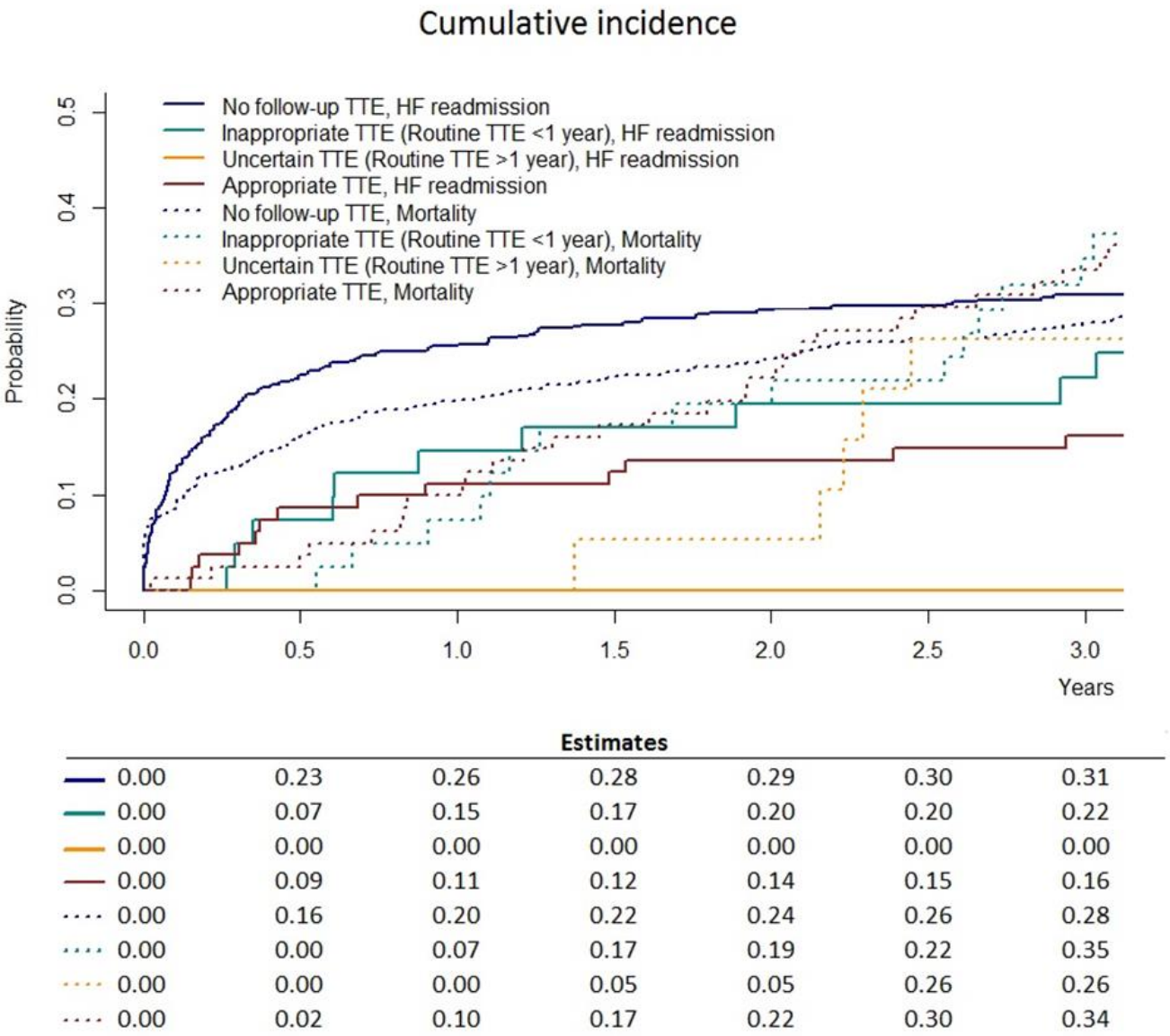
Risk of death 3 years	2.31 (1.12 - 4.78)	0.02			7.60 (3.79 - 15.20)	<0.01		
Betablockers use	1.05 (0.79 - 1.40)	0.73			0.68 (0.53 - 0.88)	<0.01		
Diuretics use	1.71 (0.99 - 2.96)	0.05			0.66 (0.46 - 0.95)	0.02		
Mineralocorticoids	1.14 (0.85 - 1.53)	0.39			0.97 (0.74 - 1.27)	0.81		
Angiotensin-converting-enzyme inhibitor/ angiotensin-II receptor blockers	1.21 (0.86 - 1.70)	0.28			0.66 (0.50 - 0.86)	<0.01	0.68 (0.52 - 0.89)	<0.01
Calcium-antagonists	1.27 (0.91 - 1.77)	0.16			0.81 (0.58 - 1.12)	0.20		
Antiarrhythmics	1.00 (0.63 - 1.57)	0.99			1.18 (0.81 - 1.73)	0.39		
Digoxin	1.10 (0.79 - 1.55)	0.57			0.97 (0.71 - 1.33)	0.87		
Statins	1.50 (1.12 – 2.00)	0.01	1.41 (1.04 - 1.90)	0.03	0.67 (0.52 - 0.87)	<0.01		
Hypertension	1.16 (0.84 - 1.61)	0.36			0.92 (0.70 - 1.21)	0.55		
Dyslipidemia	1.27 (0.96 - 1.68)	0.10			0.85 (0.66 - 1.09)	0.20		
History of angina	1.17 (0.88 - 1.56)	0.29			0.71 (0.53 - 0.94)	0.02	0.62 (0.46 - 0.83)	<0.01
History of atrial fibrillation	0.99 (0.75 - 1.32)	0.97			1.17 (0.90 - 1.5)	0.24		
Arrhythmia	0.77 (0.51 - 1.14)	0.19			1.36 (1.01 - 1.82)	0.04	1.46 (1.07 - 1.98)	0.02
Cardiomyopathy	1.05 (0.79 - 1.41)	0.73			1.00 (0.77 - 1.29)	0.99		
Dilated cardiomyopathy	1.16 (0.85 - 1.58)	0.36			0.87 (0.66 - 1.15)	0.33		

Hypertrophic cardiomyopathy	0.71 (0.37 - 1.39)	0.32			1.50 (0.94 - 2.38)	0.09		
Restrictive cardiomyopathy	1.00 (0.43 - 2.29)	0.99			1.07 (0.47 - 2.41)	0.88		
Deep vein thrombosis	0.94 (0.47 - 1.90)	0.87			1.23 (0.64 - 2.36)	0.53		
Angioplasty	0.99 (0.7 0- 1.39)	0.94			0.74 (0.53 - 1.05)	0.09		
Myocardial infarction	1.23 (0.93 - 1.63)	0.15			1.07 (0.82 - 1.38)	0.63	1.35 (1.02 - 1.79)	0.04
Cerebrovascular disease	0.96 (0.68 - 1.36)	0.83			1.24 (0.93 - 1.65)	0.15		
Renal disease	1.82 (1.35 - 2.44)	<0.01	1.76 (1.31 - 2.37)	<0.01	0.88 (0.65 - 1.20)	0.42		
Valvular disease	1.32 (0.99 - 1.75)	0.06			1.01 (0.78 - 1.31)	0.93		
Cardiac catheterisation	1.75 (1.23 - 2.50)	<0.01	1.68 (1.18 - 2.41)	<0.01	0.39 (0.22 - 0.67)	<0.01	0.40 (0.23 - 0.71)	<0.01
Diabetes mellitus	1.04 (0.78 - 1.38)	0.79			0.84 (0.65 - 1.09)	0.20		
Chronic obstructive pulmonary disease	1.07 (0.81 - 1.42)	0.65			1.10 (0.85 - 1.42)	0.46		
NHYA class 3 and 4	1.34 (1.01 - 1.77)	0.04	1.42 (1.06 - 1.90)	0.02	1.51 (1.17 - 1.95)	<0.01	1.34 (1.04 - 1.74)	0.02
Active change in management after follow-up TTE	0.81 (0.50 - 1.31)	0.39			0.80 (0.43 - 1.50)	0.49		

Cumulative incidence function for HF readmission and mortality.

Cumulative incidence curves for HF readmission and mortality for all TTE groups are shown in Figure 6- 3. The curves for HF readmission ($p<0.01$) but not mortality ($p=0.61$) showed differences between groups.

Figure 6- 3 Cumulative incidence curves for HF readmission and death between TTE groups



Estimates						
0.00	0.23	0.26	0.28	0.29	0.30	0.31
0.00	0.07	0.15	0.17	0.20	0.20	0.22
0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.09	0.11	0.12	0.14	0.15	0.16
0.00	0.16	0.20	0.22	0.24	0.26	0.28
0.00	0.00	0.07	0.17	0.19	0.22	0.35
0.00	0.00	0.00	0.05	0.05	0.26	0.26
0.00	0.02	0.10	0.17	0.22	0.30	0.34

Impact of change in management after follow-up TTE.

Table 6- 5 shows the competing risk regression for TTE groups and its active change in management derived. No statistically significant differences were found between groups of TTE on HF readmission or mortality neither a statistically significant impact of active change in management (HR for readmission after change in management in the presence of death as the competing event: 1.17 (95% CI: 0.5, 2.75), p=0.72) (Table 6- 5).

Table 6- 5 Impact of active change in management after TTE on HF readmission in the presence of death as the competing event

	Multivariable HF readmission		**Multivariable mortality**	
	HR (95% CI)	p	HR (95% CI)	p
Inappropriate	Ref	_	Ref	_
Uncertain	0.20 (0.03 - 1.44)	0.11	0.86 (0.39 - 1.89)	0.71
Appropriate	0.82 (0.38 - 1.78)	0.61	1.00 (0.62 - 1.98)	0.74
Active change in management	1.17 (0.50 - 2.75)	0.72	0.70 (0.37 - 1.32)	0.27

Discussion

To our knowledge, this is the first study that analyses survival time and HF readmission time of "appropriate", "inappropriate" and "uncertain" TTE in HF patients and makes a comparison between those tests, to understand the association of AUC status and outcome.

The results of this study support the perception that "inappropriate" tests (routine follow-up TTE <1 year) do not provide value to the time of survival or event-free time, even in the presence of active change in management. Moreover, these results also suggest that routine follow-up of stable HF patients after >1 year ("uncertain" tests) does not improve the composite of survival and HF readmission time, compared with follow-up <1 year ("inappropriate"). Even when HF readmission time is analysed in the presence of death as a competing risk, the timing

of routine follow-up TTE did not influence the outcome. The use of echocardiography according to the time expressed by the AUC, did not impact time to readmission and death.

Transthoracic echocardiography in Heart Failure.

TTE is relevant in evaluating patients with HF to clarify prognosis, evolution of disease and to determine optimal medical therapy [162]. The AUC propose that follow-up TTE in stable patients with HF after an interval of <1 year is inappropriate, but the expert panel could not reach a consensus in the setting of follow-up for >1 year. Interestingly, although previous evidence shows that TTE improves outcomes in HF patients (compared to no use of TTE), our results do not show an impact of echocardiography on free-event time or survival time. The results of the present study suggest that, although there is an increase in risk of HF readmission when there is no routine follow-up TTE, the routine evaluation with TTE in stable HF patients does not have a specific impact in risk of death or time to death.

AUC and change in management.

One of the most important roles of cardiac imaging is to guide practitioners and specialists to improve health outcomes. This aim is achieved by choosing the correct treatment for the patient. A change in management after a follow-up TTE should be the result of a detailed analysis of the patient's situation and should seek to improve pre-existing conditions. Recent data regarding downstream testing have challenged the concept of "inappropriateness" by contextualising it in relation to clinical impact [35]. These analyses, which showed that "rarely appropriate" tests may impact the management of patients, are confirmed by our study, showing that 24% of "rarely appropriate" TTE led to an active change in management. However, changes in patient management were not associated with a tangible improvement in survival time. Nonetheless, it is hard to measure the clinical impact of an echocardiogram; medical decisions are complex, and other variables readily confound the effects attributable to imaging.

The need of "rarely appropriate" testing.

Although the AUC have been used for several years and the appropriateness of cardiac imaging has been a major budgetary issue [163, 164], the rate of "inappropriate" testing has not reached zero [6, 35, 165]. Although various strategies have been used to decrease the 'rarely appropriate' ("inappropriate") tests, such as educational campaigns, feedback, audits, point of order

software and decision supports and radiology benefit managers, the results have not been uniform [99, 120, 166, 167].

The presence of "inappropriate" tests in different hospitals, communities and environments, raise the prospect of some value in these tests [35, 59, 75]. The concept of "inappropriateness" subjectively used by the AUC determines that such tests are a waste of resources that put the patients in more risk than benefit, as these patients are likely to have less pathology than those who have "appropriate" TTEs. The results of this study add another perspective: although it has previously been proven that "inappropriate" tests have led to a change in management, we cannot separate the close relationship between change in management and improvement in survival and event-free time. More research should be done to understand the change in management in a stable HF patient after an "inappropriate" test with no impact in survival.

Limitations.

These analyses have inherent limitations. First, there may be been situations where patients attended other institutions which are not part of the information system used by our hospital system. However, it is uncommon for patients to move between public and private systems from one admission to the next.

Second, although "inappropriate" tests led to a change in management, our study did not evaluate other factors such as improvement in quality of life that could be the drivers of the change of treatment after testing. Nonetheless, our study found that the change in treatment for stable HF after "inappropriate" or "uncertain" testing did not have an impact in survival time or free-event time between groups.

Third, the role of type B natriuretic peptide in guiding therapy is controversial, and this assay was performed mainly in symptomatic patients. Statistical methods for dealing with missing BNP data were not performed because these were not missing at random. Nonetheless, various biomarker approaches for assessment of volume status – including bioimpedance vector analysis [168] - may offer an important adjunct or alternative to echocardiography.

Finally, a multidisciplinary approach is clearly critical in the management of patients with heart failure [169]. However, to the extent that surveillance echocardiography is used in making judgments about HF management, the evidence here suggests that this has limited benefit.

Conclusion

Among patients with stable heart failure who underwent routine follow-up TTE, the timing of this test (or indeed its performance) did not influence survival or HF readmission time. These findings suggest that the distinction of appropriateness of routine repeat TTE in stable HF, based on the timing of testing after index admission, appears unjustified.

Chapter 7. Understanding Cardiac Imaging Decision-Making: Appropriate Use Determinants

Preface

The previous chapters have demonstrated a need to improve the appropriate use of imaging in Australia in order to impact health costs and patient care.

In consideration of using the AUC in Australia in support of the above, inconsistencies between guidelines and criteria have been identified, as well as the limited impact that the introduction of AUC has had on ordering behaviour and health outcomes.

To better understand why behaviours and outcomes have not been affected following the publication of AUC, the following study was completed in an attempt to determine a relationship between appropriate use and decision-making.

Abstract

Background: The Appropriateness Criteria (AUC) for Echocardiography were designed to facilitate doctors' decision-making, to reduce variability, and to achieve appropriate use. However, there is little evidence that the AUC have had a sustained impact on clinician's ordering behaviour. This study explores the professional, systemic, policy and patient-related factors that contribute to current echocardiography ordering behaviour, with a focus on the appropriate use of echocardiography.

Methods: Semi-structured face-to-face interviews with cardiologists and non-cardiologists who had requested echocardiograms at a tertiary hospital were conducted. The interview guide included hypothetical clinical scenarios to better understand doctors' decision-making in ordering echocardiograms and the actions they could take when receiving test reports. The methods of analysis were a combination of thematic and discourse analysis.

Results: 17 clinicians were interviewed, 10 of them were cardiologists. The primary reason for requesting an echocardiogram was a medical reason. However, other factors influenced decisions. Three main categories emerged from the interviews: personal factors, systemic factors, and follow-up of guidelines and protocols. Personal factors included training experience, attitude towards workload, the perception of risks, relationship with colleagues, patients' expectations, and management of uncertainty. Systemic factors involved the availability of services and patients' information.

Conclusion: The idea of appropriate use is a subjective concept, which is related to the need of performing a test which is hard to capture in AUC. This difference in concepts plus the variety of factors that influence decision-making makes the AUC ill-equipped to achieve appropriate use of echocardiography. The drivers of inappropriate testing are not entirely controlled by AUC.

Introduction

In Australia, data have shown a consistent increase in the use of echocardiography during the last decade. The marked geographical differences in its use support the perception that the utilisation of cardiac imaging is not necessarily illness-related and its use should be framed under the premises of the appropriate use [4]. In the United States of America, cardiovascular imaging procedures are commonly listed within the first 200 Medicare expenditures, echocardiography being the most used cardiac imaging test [18].

Guidelines for the use of cardiac ultrasound, including the Appropriate Use Criteria (AUC) for Echocardiography, have been developed to help clinicians to achieve proper use to control overuse of imaging services and to improve patient care [26, 65]. Despite the AUC, it has been estimated that 5 to 25% of cardiac imaging examinations are still being performed for "inappropriate" indications [6, 170].

There have been several strategies in the USA promoting and enforcing the systemic use of the Appropriate Use Criteria to effect a decrease in the percentage of "inappropriate" tests [120-122, 170-173]. Such approaches include educational interventions, clinician audit and feedback, use of radiology benefits management companies and pre-authorisation processes, and use of software at the point of request [99, 120-122, 170-173]. However, the impact that those interventions have had on clinicians' ordering behaviour to adhere to the Appropriate Use Criteria has been questioned. Although there has been a decline in cardiac imaging in the USA in the last decade, it has not applied to all imaging techniques [25, 67] . Moreover, studies suggest that there has been limited impact on the improvement of resource utilisation and appropriateness over time [6, 121].

To address the issues associated with decisions around the appropriate use of echocardiography, it is, therefore, necessary to examine the factors that lead a doctor to choose when to perform a cardiac test and most importantly, to explore the aspects that influence doctors to order an echocardiogram that is "inappropriate". Because requesting echocardiograms is not exclusive to cardiologists [6, 174], the need for addressing those factors in other specialities is mandatory.

This chapter is a qualitative study aimed at understanding reasons for ordering imaging tests and factors contributing to doctors (cardiologists and non-cardiologists), prescribing inappropriate echocardiograms. The process of decision making regarding cardiac imaging, particularly echocardiography, is explored, including the impact of available resources, consideration of potential benefits, harms and other factors influencing decision making.

Understanding these factors may assist in the development of more effective recommendations and educational campaigns to reduce the incidence of, and ultimately avoid, inappropriate testing.

Methodology

Study design

We conducted a qualitative research study using individual face-to-face semi-structured interviews that included doctors (cardiologists and non-cardiologists) who had ordered an echocardiogram during the past year at the Royal Hobart Hospital (RHH). The RHH is Tasmania's largest hospital and primary referral centre. The Tasmanian Ethics Committee approved this study.

Participants were selected and interviewed between April and September 2016. We guided the interviews by a set of open-ended questions developed through literature review and expert consultation. Doctors were asked to provide some general information about themselves and the characteristics of their practice. The interview guide included hypothetical clinical scenarios to understand better doctors' decision-making on ordering echocardiograms and the actions they could take when receiving test reports. The cases included "appropriate" and "inappropriate" indications according to the Appropriate Use Criteria for Echocardiography[26]. Additionally, the scenarios included a patient with an "inappropriate" echocardiogram with incidental findings to uncover concerns, perspectives, and actions about receiving unexpected abnormal results. Interviews focused on the factors that influence the decision of clinicians to order an echocardiogram, including the role patient expectations play in the decision-making process. We also examined their awareness of Appropriate Use Criteria, how discussions about performing an echocardiogram were initiated with patients, and what they understood their role was in avoiding inappropriate requests. The process of decision making about cardiac imaging was explored, including the impact of available resources, consideration of potential benefits and harms and other factors required for decision making. Interview questions and clinical scenarios are provided in Appendix 6- 1 and Appendix 6- 2.

The methods of analysis were a combination of thematic and discourse analysis[175-177]. The thematic analysis allowed the identification of common factors that shape the experience of healthcare professionals and their patients. Discourse analysis focused on the structural and

cultural factors that shape individual experiences, factors that may not always be at the forefront of interviewees' awareness.

Data collection

Potential participants were contacted by telephone, email, and in person, and asked if they were interested in being involved in the study. Cardiologists were chosen from the Cardiology Department and non-cardiologists from the General Medicine Department. These two departments originate the vast majority of echocardiogram requests at the hospital. An approach document describing the study and inviting them to participate in this research (Appendix 6- 3 and Appendix 6- 4), were sent to the clinicians (Cardiologists, Cardiac Surgeons, General Medicine Specialists, Pneumologists, Nephrologists, Neurologists, Gastroenterologists, Geriatricians, Registrars for each speciality, and Interns). For doctors who chose to be interviewed, a written informed consent was completed prior to the meeting (Appendix 6- 5).

Physicians were interviewed in person. The interviews were designed to last between 25 and 30 minutes. All participants gave verbal and written consent. No reward was given after the interviews.

The face-to-face interviews were entirely recorded and transcribed verbatim.

Each participant was identified by a unique numerical code (individual ID code). Recordings, transcriptions and interviewees details were stored electronically using the individual ID code for each participant. Consent forms were stored separately. All information was used and presented in a deidentified manner.

Data analysis

Thematic codes were established using two ways: *a priori* (found by literature review and interview guide) and through the reading of the transcripts. Each code was clearly defined to assure accuracy. One researcher (RF with less than a year in qualitative research) coded all interviews using Excel®. Coding was closely supervised by a qualitative researcher-author (KJ with more than 5 years of qualitative research). Discrepancies in coding were resolved by consensus.

Results

General characteristics of participants

Twenty-seven clinicians were invited to participate (Table 7- 1). Seventeen physicians (63% of the invitees) participated in the study and divided into two separate groups: "cardiologists" and "non-cardiologists". The rate of participation was higher in the "cardiologists" group (77% vs. 50% in non-cardiologists).

The "cardiology group" (or cardiologists) was formed by cardiac specialists, cardiac surgeons, and registrars of such specialities. Ten of the participants were included in the "cardiology group". The "non-cardiologists' group" included: physicians who were not cardiac related specialists or cardiac related registrars and interns. Seven of all participants were included in the "non-cardiology group". These physicians were General Medicine Specialists, Neurologists, Geriatricians, and interns (doctors who had been practising within the first year after finishing medical school). In total, fifteen of the clinicians in the group were specialists or speciality registrars; all of them had been practising for five (or more) years since graduating from medical school. The other two physicians were interns.

Table 7- 1 . Rate of response to invitation to participate in the study

	Contacted	**Participated**	**%**
Cardiologist (<5 years of medical experience)	0	0	-
Cardiologist (5-10 years of medical experience)	6	5	83%
Cardiologist >10 years of medical experience	7	5	71%
Non-cardiologist <5 years of medical experience	2	2	100%
Non-cardiologist 5-10 years fo medical experience	6	2	33%
Non-cardiologist >10 years fo medical experience	6	3	50%
Total	27	17	63%

Gender diversity was skewed towards males in the cardiology group due to there being no female cardiologists currently working in the hospital. Cardiologists were represented by the highest proportion of specialists having practiced for more than a decade (4 out of 10 vs. 2 out of 7 in the non-cardiology group).

There was a marked difference in the number of echocardiography requests per month between groups as expected (Table 7- 2).

Table 7- 2. Characteristics of participating clinicians (n=17)

	Cardiologist	**Non-cardiologist**
n	10	7
Age years median (range)	38 (29 – 70)	36 (26 – 41)
Gender, Male (%)	10 (100.0)	4 (57.1)
Years since Medical School , median (range)	10 (5 – 48)	8 (0.9 – 15)
Cardiology patients per month (mean (sd))	189 (128)	49 (48)
Echocardiography requests per month, median (range)	20 (8 – 100)	4 (1 – 20)

Quotes

Summative testimonials for the main themes were supported by quotations. Quotations were represented as "NC" or "C" if the participant was in the "Non-cardiologist" or in the "Cardiologist" group respectively. Identification of clinicians was avoided by dividing them according to their years of clinical practice: less than 5 years of clinical practice, between 5 and 10 years of clinical practice, and more than 10 years of clinical practice. Years of experience as specialists was not chosen to characterise those groups because some of the interviewees were not specialists (registrars in speciality training or interns).

Clinical reasoning process and Appropriate use concept.

All of the participants reported that the primary motive for ordering an echocardiogram was a medical reason, either to improve existing comorbidities or to avoid cardiac-related morbidity in the future. This medical reason involves not purely objective clinical assessment (i.e. symptoms or signs, age, body characteristics of the patient), but also factored a consideration of the patients' quality of life:

"If you have a 90-year-old patient with dementia who's had a stroke, where the care is going to be palliative, in that situation you would never order an echo as you know it's not going to be useful." (NC5, more than ten years of clinical practice).

Towards those aims, all participants reported that an echocardiogram was a valuable test regardless its limitations. Despite overall positive influence that echocardiography has in their practice, they were aware of the impact that overuse of health services has on the health system and the health budget. All the interviewees agreed that an appropriate use of health services, and in this case, appropriate use of echocardiography, is necessary.

When the participants defined "appropriate use", they indicated that an "appropriate" echocardiogram is a test that helps them to make a diagnosis and to plan further treatment. This understanding was not the result of knowledge or awareness of the Appropriate Use Criteria for Echocardiography. Non-cardiologists expressed that they did not know about the existence of the Appropriate Use Criteria for Echocardiography, and, although all cardiologists said they were aware of the AUC, four cardiologists acknowledged they had read the Criteria, and one cardiologist stated that he tries to use it in the everyday practice. The interviewees used the terms "useful" and "necessary" rather than "appropriate" (or "inappropriate") when discussing the value of echocardiograms.

However, the concept of "appropriateness" is a personal understanding of the conditions of the patient and circumstances that involves the doctor-patient encounter, that leads doctors to order or perform an echocardiogram. They believe that this "subjectiveness" can lead to misinterpretations:

"It is easy for other people to misunderstand a test as inappropriate when retrospective evaluations are done; they won't have the full context of the assessment that the doctors did at that moment." (C1, 5-10 years of clinical experience).

Factors that influence clinical reasoning and echocardiography ordering.

Coding of transcripts identified that there were three major categories of factors that influenced physicians for requesting "appropriate" echocardiograms: (a) "Personal" factors, (b) "Systemic" factors and (c) "Guidelines and protocols" (Figure 7- 1).

Within these three major categories, some themes emerged and were collapsed in the following nine themes:

(a) Personal factors included:

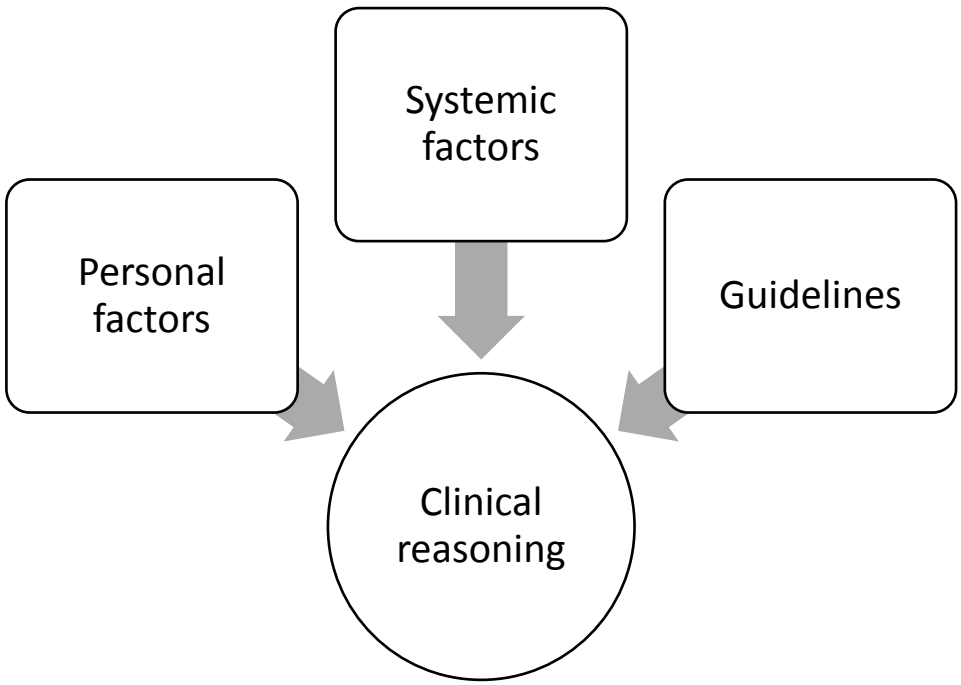
- Training and medical experience,
- Attitude towards workload
- Perception of risks and harms,
- Relationship with colleagues,
- Patient's expectations,
- Management of uncertainty.

(b) Systemic factors included

- availability of services
- availability of patient's medical information

(c) Guidelines and protocols included the follow-up of guidelines, rules and protocols.

Figure 7- 1 Factors that influence clinical reasoning and echocardiography ordering.



Personal factors.

Training and medical experience

Doctors develop expertise, competence and medical discernment through years of practice, academic and practical training, and previous case experience.

Participants commonly reported that the improvement of their skills and development of expertise were the most important elements leading to ordering an echocardiogram, impacting their rationale and rationalisation for requesting exams. Senior cardiologists indicated that these skills help them to understand when to perform an echocardiogram and in what sort of patients. Doctors with fewer years of practice noted that the time for follow-up is a challenge.

"You develop a technique that I call 'the pattern recognition'. You understand the natural cause of disease and the point in time at which you have to make a decision to do something different; it is based upon clinical assessment, not upon an echocardiogram" (C7, more than 10 years of clinical practice).

However, doctors with more years of medical practice indicated that, although they welcomed the improvement in echo techniques and availability of exams, differences in the training process of new doctors was a possible cause of overuse and "inappropriate" use of echocardiography:

"There is less emphasis on clinical examination nowadays than it used to be in the past. Nowadays we hear murmurs, and we try to quantify their severity which leads straight to ordering an echo to get the actual accurate quantity of information. However, this can result in overuse of imaging." (C8, more than 10 years of medical practice).

Along with the change in the training process and the broad availability of echocardiography, came a challenge that doctors, particularly cardiologists, stated: a better understanding of the technical issues of echocardiography and the scope and limitations of the test when evaluating cardiac diseases:

"I think there is a bit of a misconception amongst non-cardiologist; They do not understand what information we get from an echo, they do not know that we cannot actually explain everything from an echo. They are always obsessed with the ejection fraction, but other information and findings are of concern."(C2, 5-10 years of medical practice).

It is common for a junior physician to defer to a senior consultant or specialist in respect to ordering tests due to a lack of clinical experience. Additionally, non-cardiologists follow the suggestions of the cardiac specialists because they recognise the experience and specialised skills and knowledge.

Attitude towards workload

Some doctors acknowledged that the ordering behaviour can be impacted by their workload or mood, resulting in an increase of "unnecessary" requests:

"Sometimes it is just laziness. If you are very busy on the ward and it is too hard to decide whether it is entirely necessary or not, it is just easy to get one, and if it is unnecessary, no one usually says anything, but if they wanted it and it is not there, then it becomes an issue." (C4, 5-10 years of medical practice).

"Some days patients want tests that I feel are not necessary but I want to avoid discussions or I'm tired and I will order tests anyway as I can't be bothered trying to change the patient's expectations". (NC3, more than 10 years of medical practice)

Perception of risks and harms

The doctor's understanding about limited hazards, harms, and possible misdiagnosis does not impact ordering patterns for echocardiograms. Although all doctors noted the possibility of risks, harms, and/or misdiagnosis with transthoracic echocardiography, it was considered safe to perform because it does not involve radiation or invasiveness. Clinicians confirm that risks increase when contrast is used, but this does not impact their decisions to request tests. All of them stated that they usually order echocardiograms because there is a problem they want to solve, and therefore, some sorts of findings are "expected". Of note is the lack of concern regarding the risk of malpractice when considering requesting tests in the Australian environment.

Relationship with colleagues

Cardiologist, particularly, may request exams if peers believe it is necessary, even when their thoughts differ. This attitude decreases the risk of bad relationships with colleagues:

"When I was a trainee, I used to require echoes because I received orders from superiors who had the medical-legal responsibility. However, being a more senior doctor, I don't worry

about 'getting in trouble' if I don't order an echo, but it is just to keep the clinical process running smoothly." (C1, 5-10 years of medical experience).

"If an experienced cardiology colleague says we should do another echo, I would not feel strong enough to say no. I think to myself 'what a waste of time', but I would not say it. You know, people have different views, sometimes they are wrong, but they might think 'I am not'." (C7, more than 10 years of medical practice).

Patient's expectations

All doctors with more than 5 years of medical practice, stated they feel an expectation is placed on them by patients, that a test should be done, especially in an ambulatory setting, and believe that skills gained through years of clinical practice have given them the ability to manage these effectively.

These physicians believe three main reasons increase the probability of patients to have expectations: the available information predominantly on the internet, the level of a patient's private insurance cover, and the need to be referred to a cardiac specialist. Cardiologists are more likely to request echocardiograms in the latter scenario.

Apart from a patient's expectation, the anxiety of some patients may lead to an echocardiography request. In this situation, the test is seen as a way to assuage the patient and avoid new consultations in the future:

"Patients want reassurance with a test, and I think it helps in the long run because chances are it will be normal, and it will help them realise that or at least reassure them that what they are feeling, 'the funny pains', are not related to pathology. It can reduce the anxiety and prevent representations to the hospital, helping to keep them from coming in with chest pains so it may be appropriate for that purpose." (C4, 5-10 years of medical practice).

Uncertainty

Despite the experience that doctors gain during their years of practice, uncertainty plays a significant role in the decision-making process, even in very experienced specialists:

"I am worried if they don't have a full assessment and I miss something that it is going on with their heart that is not apparent because ECGs and clinical examinations are not very precise."(C9, more than 10 years of medical practice).

"There is a variation in how we manage uncertainty and anxiety. Some doctors manage that uncertainty better than others. I have seen consultants who are very anxious not understanding what is happening to the patient, and they request more tests." (NC3, more than 10 years of medical practice)

This management of anxiety and uncertainty may be caused by previous clinical case experience, and it increases the chance of "inappropriate" requests:

"There might be a bias to a situation where they missed an important finding, when they were a junior doctor, so they always do scans because they are worried that something might happen like years ago." (NC3, more than 10 years of medical practice)

Systemic factors

Availability of services

All respondents noted that the ease of access to echocardiography can impact decisions to order tests. Barriers, such as distance to the point of service or requirements needed to enable patients to move freely also affect the ordering of tests (e.g. a patient in a wheelchair). However, the availability of specialised transport services can assist to address these barriers.

In the hospital, the waiting time until the test can be conducted was an important consideration for non-cardiologists because of the pressure they have to discharge patients.

Cardiologists identified that the private setting increases the likelihood of ordering echocardiograms because the echocardiogram is more accessible regarding waiting time and there are economic incentives for both, hospital/practices and specialists that report the exams.

Availability of patient's medical information

One of the issues that the clinicians postulated was that the system does not readily provide them with all the mandatory information for a patient's consultation; the poor availability of previous echocardiogram results can create the environment for duplicate examinations, even though the earlier test may have been done a short time prior.

Protocols and guidelines

Although doctors stated that guidelines and protocols are useful to standardise practice and to improve health outcomes, participants also reported that they usually do not follow those documents closely. In some cases, the use of guidelines and protocols impact the ordering of

echocardiography negatively; cardiologists commented that in some scenarios, the guidelines and health policies make them request echoes, although they consider them redundant:

"There are situations where I've ordered an echo when I otherwise would not have, because guidelines mandated. In patients with pulmonary hypertension, you're required to do six-monthly echoes despite, at least in my experience, I don't find much use of those echoes. (C1, 5-10 years of medical practice).

Another physician reported that "suspicion of endocarditis" leads to ordering echocardiograms because they need to do it before requesting a transoesophageal echocardiogram which is a better test for the purpose. However, they cannot order a transoesophageal echocardiogram without a transthoracic echocardiogram previously.

Discussion

In this study, we examined factors leading doctors to order transthoracic echocardiograms and their relationship with the appropriate use of echocardiography. To our knowledge, this is the first study that intends to elucidate the thoughts and experiences of doctors when they are requesting cardiac ultrasounds to better understand the decision-making process around the appropriate use and cardiac imaging value.

The findings highlight how a variety of elements influence the decision of requesting echocardiograms in clinical practice, but most importantly, that "appropriate" use is a subjective evaluation targeting the judgements and choices the physician has to address in medical practice, rather than the result of pressures aiming to control overuse of services and sustainability of the heatlh system after following the criteria for appropriate use of echocardiography.

During the interviews, it was initially found that all doctors believed they required cardiac ultrasounds "appropriately" due to medical reasons as primary causes. However, the medical motivation was not undoubtedly a factor for "appropriate" use of echocardiography; other stimuli influenced the clinicians in some circumstances, which would lead to tests that do not have value for them, resulting in an measurable overuse and "inappropriate" use of resources. This is a significant challenge to solve when the quality of life, most commonly in the elderly or people with multiple co-morbidities or sequelae, is assessed, or when clinicians have no clarity of when to re-assess the patient's case (e.g. timing for follow-up).

One of the most remarkable findings of this study is that there is a dissimilarity between the concept of appropriateness within the Appropriate Use Criteria [26] and that held by doctors. For the Appropriate Use Criteria, the "negative consequences" of the test which include procedural risks, false negatives (or false positives), or misdiagnosis, are the pillar for the definition of "appropriate" use of transthoracic cardiac ultrasound. While the risk of undesirable consequences or the possibility of diagnosis errors with "inappropriate" tests were certain drivers for the development of the AUC [26], the clinicians in the current investigation stated that, while these risks might be acknowledged, are not the primary driver for not ordering an examination. The concept of "appropriate use" is influenced by the "necessity" of the test, which is in fact, the main reason for requesting an echocardiogram to lessen the uncertainties that they are facing with the patient. Moreover, despite awareness about the close relationship between health expenditure and use of medical services, this reason was not the motive for "appropriate" use of echocardiography either. Economic reasons come to play a significant role only when the patient is not covered by health insurance or needed to pay copayments[178].

The "necessity" and "appropriateness" of a test are comparable according to the interviewed physicians. The most valuable, and therefore, appropriate test for them, is a test which is needed to make decisions or to impact in some way, the care that the patient is receiving. The most significant factors influencing doctors to order "valuable" and "appropriate" tests are their experience and medical expertise. Doctors with more years of experience, more training, or with particular previous medical experiences, believe that they understand the diseases and can better rationalise the use of health resources, including echocardiograms. This situation is enhanced with better understanding of the limitations of the echocardiography technique, which is more common among cardiologists and related specialities. However, it is critical to take into account that the management of uncertainty[178, 179] is an independent factor, that is not completely mitigated with more years of experience or training. There were differences between management of uncertainty even within the most experienced cardiac specialists.

Professional experience helps doctors to manage patients' expectations. However, there are some situations where the patient's expectations play a vital role in the decision of requesting exams. In this environment, accepting patient's expectations, although contrary to physicians' desires, is not always seen as a negative factor, because it may help to improve the doctor-patient relationship and most importantly, to impact positively patient's treatment by decreasing patient's anxiety, which usually interferes with better outcomes. In addition,

requesting echocardiograms to decrease patients' anxiety, is thought to decrease the chance of patient re-visits to hospitals and clinics. For these physicians, reducing patients' anxiety can be one of the most "appropriate" echocardiography indications. Interestingly, although professional experience helps doctors to interact better with other clinicians and colleagues, requesting an echocardiogram because a colleague with similar experience, or superior, has requested it, it is usually not seen as "appropriate". In this scenario, doctors order the exams to avoid any sort of future controversies among clinicians, but the echocardiogram is still seen as "inappropriate".

Other sorts of factors were discussed during the interviews which impact the decision to request echocardiograms. Easy access to an exam is an important factor. Barriers to easy access (e.g long waiting lists due to few echocardiography laboratories) may result in the underuse or constrained ordering of tests as has been described[180]. However, the decision-making process when there are barriers to test access, results in a more "appropriate" use of resources: the tests are requested if there is a real need. This was mainly identified among non-cardiologists whose work is predominantly with inpatients. Long waiting periods to perfrom echocardiograms in hospital were a significant consideration, because it is ideal for patients to be discharged once they are stabilised and the management for the outpatient setting organised. In the outpatient setting, these barriers can be avoided if the patient has a private insurance. Moreover to easy access, the private setting may lead to more testing because of the existence of some economic incentives (for the specialist who interprets the results or the echo laboratory), a situation that has been found as a cause of overtesting by other authors [178].

However, easy access to testing when a patient has a private insurance is not the only reason for requesting "inappropriate" echocardiograms related to the availability of resources. Doctors expressed that lack of availability of previous results are an important cause of re-testing. Medical practices do not share the same digital record system and commonly, when tests are performed, the results are not immediately available. Generally, the quality of the tests is not a problem: doctors trust the professionals who perform echocardiograms and the physicians who interpret the exam. However, quality may be an issue when patients have echocardiography results from other echocardiography laboratories.

Other factors such as following guidelines and protocols, encourage doctors to request tests which, sometimes, are not seen as appropriate because the management plan is usually scheduled before the results of the test. This is particularly common in patients with pulmonary

hypertension, who have a follow-up echocardiogram every six months to maintain eligibility for treatment as a part of prescribing policy, or patients with suspected endocarditis who need a transthoracic echocardiogram to access to a transoesophageal exam.

How to improve appropriate use

The subjectiveness of the concept of appropriate use among clinicians and the variety of factors that influence their decision-making, do not favour the use of the appropriateness criteria as the way to accomplish appropriate use of echocardiography. Appropriate use of cardiac imaging is a hard objective to achieve, which involves not only doctors but also patients and policy makers.

Academic training in cardiovascular diseases and echocardiography, health economics, and educational campaigns in general, including feedback and audit processes, would help physicians to understand the natural development of diseases and to find the appropriate time to follow-up a patient, and therefore, to order a test when it is necessary. These approaches may be enhanced by the use of software at the point of request (to give information to doctors about the benefits and harms of performing tests), and availability of staff at the echo-laboratory laboratory to pick up possible inappropriate situations, including screening imaging requests. Moreover, patients may play a vital role in the proper use of echocardiography by actively interacting with the physician through shared decision-making and health literacy. These approaches would certainly help physicians in their management decisions, including management of uncertainty.

Policy makers and governments have an important role in the appropriate use process. Imbalance between access to healthcare between public and private settings, financial incentives, and systems of payment influence directly the decision making of doctors, resulting in inappropriate use of resources. Interaction between all these actors and across all levels of care impact the quality of care and appropriate use of resources.

Strengths and limitations

Although we achieved saturation, as with any other qualitative study, there are limitations with the present study and the results are not applicable to all medical environments. The purpose of this study was to investigate the challenges that doctors face when ordering tests to understand how to achieve appropriate use. Potential limitations of this research are its small size, and the restricted response to participate in the study particularly in the “non-cardiology”

group. Since the participants were recruited from one hospital setting, this could be considered a limitation in terms of transferability. However, to limit the potential weakness of the design of the study, the clinicians interviewed during this investigation varied in terms of clinical expertise and experience in the inpatient and outpatient settings, and the use of hypoteical clinical scenarios included inpatient and ambulatory patients, allow the understanding of more factors that influence professionals when requesting echocardiograms.

To decrease the possible bias, an experienced researcher was included in the analysis of codes and the grouping of those into themes, and the final allocation to themes was attained by consensus among all investigators.

Conclusion

Thhis research found that the drivers of inappropriate testing are not all controlable by AUC. The results suggest a mistmatch between the clinical reasoning process of physicians and the AUC for echocardiography.

Chapter 8. Identification of inappropriate echocardiogram requests at the point-of-service

The research contained within this chapter has been published as[8, 181]:

- Fonseca R, Pathan F and Marwick T. Identification of Rarely Appropriate Echocardiograms in the Echocardiography Laboratory. Journal of the American College of Cardiology. 2016;67:1603-1603.*
- Fonseca R, Pathan F and Marwick TH. Development and validation of a screening tool for the identification of inappropriate transthoracic echocardiograms. BMJ Open. 2016;6:e012702.*

**Drs Fonseca and Pathan should be regarded as joint first authors.*

Preface

The chapters 4 to 7 have analysed several potential issues in the adoption of the AUC in Australian practice. These problems include differences between guidelines and criteria, no clear impact on clinicians' requesting behaviour, unjustified definition of inappropriate indications based on follow-up time in patients with heart failure, and some disparity between the clinical reasoning process of doctors and the AUC. The problems may cause limited acceptance of the AUC for use in Australian practice. Irrespective of these challenges, it is evident, as demonstrated in chapter 2, there is a need to introduce a method or methods to ensure appropriate use in this country, most importantly in echocardiography use.

The majority of strategies that have been implemented in an attempt to achieve appropriate use of imaging overseas have primarily been focused on the restriction of labelled inappropriate requests at the point-of-care (or point of request). The AUC model adopted in the USA aims to avoid the "inappropriate" test ordering through a system of cancellation once a test has been identified/labelled as inappropriate. However, new studies have shown that some tests currently labelled as inappropriate have in fact led to a change in management due to new important findings[35-37, 75]. Additionally, the possibility exists that requesting doctors choose "appropriate" indications instead of the patients' presenting problem in order to be seen as to adhere to the criteria (thus avoiding audits and penalties), making the development of a strategy aiming to determine potential inappropriate echo requests in a transparent way and avoid risks of

cancelling a labelled inappropriate test (due to potential of missing important cardiac findings) necessary.

For these reasons, this and the following chapter aim to explore and develop strategies to mitigate these issues. Chapter 8 seeks to determine the characteristics of scenarios labelled as "inappropriate" and to compile them in an easy-to-use questionnaire at the echo laboratory (point of service) to identify possible inappropriate requests. In addition to this strategy, chapter 8 gives us a safe approach in order to avoid inappropriate tests without the risk of missing important cardiac findings through the use of hand-held echocardiography.

Abstract

Objective:

We sought whether simple clinical markers could be used in a questionnaire for recognition of inappropriate (or Rarely Appropriate, RA) tests at point-of-service. Most applications of Appropriateness Criteria (AC) for transthoracic echocardiogram (TTE) have been at the point of order, but a simple means of identifying RA tests in an audit process would be of value.

Design, setting and participants:

The study was performed in two major hospitals in Tasmania. Two reviewers created a questionnaire based on 4 questions most commonly associated with RA (suspected endocarditis with no positive blood cultures or new murmur, lack of cardiovascular symptoms or no change in clinical status or cardiac exam, routine surveillance and previous TTE within a year) in a derivation cohort of 814 patients. This was prospectively applied to 499 TTEs to calculate sensitivity and specificity for prediction of RA, and validated in external group (n=880).

Results:

Of 499 prospective TTEs, the questionnaire selected 18% requests as being potentially RA. As 7.4% were actually RA (kappa 89%), the sensitivity and specificity of the questionnaire were 84% and 87% respectively. In the external validation cohort, the model found 11% requests needed to be screened for appropriateness with a sensitivity and specificity of 80% and 95%.

Conclusion:

A questionnaire based on 4 questions detects a high proportion of RA TTE, and could be used for audit.

Strengths and limitations:

- Four binary questions encapsulate characteristics of rarely appropriate tests according to the Appropriate Use criteria for Echocardiography.
- The questionnaire, applied to the transthoracic echocardiogram requests, selected around 1 in 5 requests as being potentially rarely appropriate.
- Two or more affirmative answers had a high sensitivity and specificity to discriminate Rarely Appropriate tests.

- It is a feasible tool, which can be used at the point of service to screen for inappropriate tests with a low impact on the workflow.
- The use of this approach requires review of medical records to adjudicate appropriateness when inadequate information is provided on the request form.

Introduction

Investigations constitute an important component of resource utilization and waste in medicine. The Appropriateness Criteria (AC) for transthoracic echocardiography (TTE) [26] have become widely adopted in the United States, and seek to control resource utilisation, reduce variability of practice, and to improve decision making and patient care [28].

Evaluations of appropriate use have exposed potential targets for improvement.[120, 182], but despite attempts to decrease rarely appropriate (RA) testing at the point of order, there is limited evidence of a decline in the number of tests, [2, 183-186] and in some cases, results are contradictory [120, 128]. AC are probably not responsible for the reduction in imaging over the last decade[85]; the trend of appropriate and RA tests has not improved during this time [6].

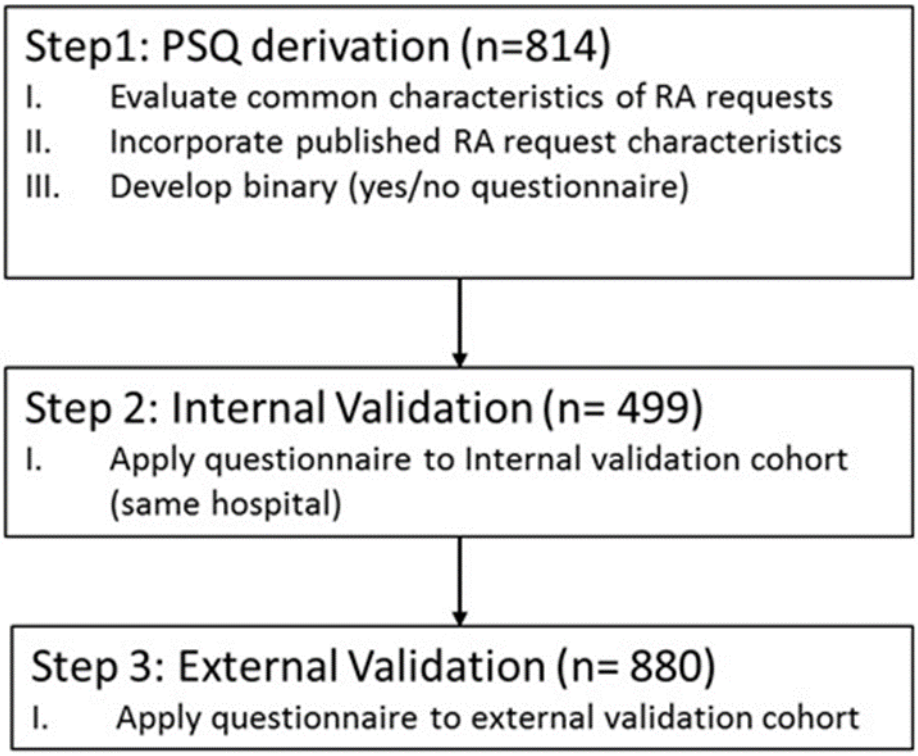
This lack of clear improvement in the rate of RA tests [6] contrasts with the heightened level of awareness on RA test use [28, 66, 183]. A recent review noted that guidelines for quality in cardiovascular imaging advocate implementation of appropriateness criteria, though there are limited effective methods to reduce the rate of inappropriate testing. They contend that part of the complexity of implementing appropriateness criteria may be unfamiliarity with the classification and the time required to review each imaging request with the guidelines[187].

Various processes have been devised to police RA requests by screening at point-of-order, for example, based on use of radiology benefit managers or software integrated with the ordering process.[121, 126, 130]. An alternative approach might be laboratory-based [2], but matching patient details against >100 AC is impractical and inefficient. In this study, we sought to determine whether a simple point-of-service-questionnaire (PSQ) based on the most common RA characteristics according to the 2011 AC for TTE [26, 59] could facilitate recognition of these tests in the echocardiography laboratory.

Methods

We sought to develop the "point of service questionnaire" (PSQ) and then perform a diagnostic accuracy analysis [188]. The model was developed in three steps (Figure 8- 1). Eligible requests were selected from two separate hospitals and at the times specified in steps 1 to 3. TTE requests were excluded for analysis if the patient was <18 years old, or if classification was not possible because of insufficient clinical documentation.

Figure 8- 1 Design of study



The study design pertained to appropriate selection of tests already ordered by the patient's physician. We elected not to discuss the uncertainties about appropriate use with these patients. Nonetheless, we have discussed appropriate use with patient representatives in meetings about medical expenditure in Australia. Because of the co-payments associated with outpatient echocardiography in this country, this is perceived as a very important topic.

Our target condition was to determine "inappropriate" (also known as "Rarely Appropriate", RA) TTE requests at the point of service as adjudged by the reference standard with the "Appropriateness Criteria for echocardiography (AC)"[26]. We developed then assessed the index test (PSQ)-

We compared the accuracy of our index PSQ model with the reference AC standard by evaluating sensitivity, specificity, positive and negative predictive value for each of the responses as well for each of the possible cut points: one affirmative answer vs two affirmative answers vs three affirmative answers.

Derivation of model:

The most common causes of inappropriate tests described by the 2011 AC for TTE were identified from a retrospective group of 814 TTE requests, at a teaching hospital. After this analysis, four questions, which summarise those characteristics, were identified. The "Point of Service Questionnaire (PSQ)" comprised binary questions based on the characteristics of RA tests in our derivation cohort and also accounted for published characteristics seen in the choosing wisely programme [187] and published research.[59].

Internal validation:

We then tested these questions in an internal validation cohort to ascertain their ability to identify RA requests as judged against the gold standard (AC code for a specific request). The four most common characteristics for RA tests (PSQ) were applied prospectively to all the requests (n= 499) for a TTE at the same tertiary referral hospital, between March and May 2015.

External validation:

The PSQ was applied to a cohort of 880 requests at a regional referral hospital between May-August 2015.

Patient demographic information, inpatient/outpatient distribution, referring physician (cardiologist or not) and the indication for the study were determined from the request form. Investigators reviewed the digital medical record to capture any additional information, especially when confronted with inadequate information. The time required to access additional information was recorded whenever such an action was necessary and the result was averaged for all such requests.

For each of the steps a general physician (RF) and a cardiologist (FP) independently recorded the results for each of the questions in the PSQ. Appropriateness of requests was scored by the same observers, independently of the PSQ evaluation, using the 2011 AC [26]. Each study was scored as appropriate, rarely appropriate (previously described as "inappropriate"), or maybe appropriate (previously "uncertain"). If the main indication was not listed in the AC, investigators were asked to select "Not classifiable".[28]. When there was disagreement, a consensus between reviewers was reached. If no consensus was attained, a third investigator (TM) reviewed the data and determined AC score. This AC score served as the "Gold Standard"

by which the PSQ was assessed. Where a repeat study was performed to guide management (e.g. repeat TTE to evaluate reverse remodelling after 3-6 months of medical therapy) despite 2 affirmative responses this request was classified as appropriate as per the AC guidelines on studies, which are used to guide management.

Sensitivity, specificity, positive likelihood ratio, and odds ratio were used to define the PSQ accuracy for the prediction of RA requests using R software[79]. The above parameters were analysed for each affirmative response and cumulatively. During individual analysis, a single affirmative response (question 1, 2, 3 or 4) was compared to "no affirmative responses". For the assessment of the cumulative affirmative responses, comparison was made to both "no affirmative responses" and to "no or one affirmative response".

Inter-observer variability in the scoring process to determine level of appropriateness was defined using kappa statistics.

Results

Predictors of RA tests were sought in a group of 814 patients among whom 9% were RA. Our results revealed the RA requests corresponded with indications where there were no new symptoms or no change in clinical status or cardiac exam (indications 35, 53, 10, 8 of the AC) and indications for routine surveillance (indications 88,11,13,40, 28). We sought to distill the underlying markers of an inappropriate request using "routine studies" and the "absence of a change in clinical status or new symptoms" which accounted for 88% of RA requests. Furthermore, 28% of RA tests had a TTE within the previous year. Additionally, there was a specific clinical situation that accounted for the 26% of inappropriate tests (Indication 53: Transient fever without evidence of bactermia or new murmur).

We identified four features associated with RA tests: evaluation in the absence of symptoms/signs of cardiac disease or no change in clinical status, routine surveillance, existence of a previous TTE within the year of the new TTE request and suspected endocarditis with no positive blood cultures or new murmur. (Table 8- 1)

Table 8- 1 . Rarely appropriate (“inappropriate”) tests found in derivation group

AC Item		Count	Proportion	Routine surveillance	Lack of change in clinical status/Evaluation of symptoms without other symptoms/signs of cardiac disease
35	Initial evaluation when there are no other symptoms or signs of valvular or structural heart disease	25	0.36	No	Yes
53	Transient fever without evidence of bacteraemia or a new murmur	18	0.26	No	Yes
10	Initial evaluation of ventricular function (e.g., screening) with no symptoms or signs of cardiovascular disease	6	0.09	No	Yes
28	Suspected pulmonary embolism in order to establish diagnosis	4	0.06	No	No
54	Transient bacteraemia with a pathogen not typically associated with infective endocarditis and/or a documented nonendovascular source of infection	4	0.06	No	No
88	Routine surveillance (<1 y) of known cardiomyopathy without a change in clinical status or cardiac exam	3	0.04	Yes	Yes
8	Light-headedness/presyncope when there are no other symptoms or signs of cardiovascular disease	2	0.03	No	Yes
11	Routine surveillance of ventricular function with known CAD and no change in clinical status or cardiac exam	2	0.03	Yes	Yes

13	Routine perioperative evaluation of ventricular function with no symptoms or signs of cardiovascular disease	2	0.03	Yes	Yes
40	Routine surveillance (<1 y) of moderate or severe valvular stenosis without a change in clinical status or cardiac exam	2	0.03	Yes	Yes
38	Routine surveillance (<3 y) of mild valvular stenosis without a change in clinical status or cardiac exam	1	0.01	Yes	Yes

Based on the above, we developed a PSQ composed of four binary questions:

Q1: Was the scan requested in the absence of new cardiovascular symptoms, or change in clinical status or cardiac examination? Note that it requires symptoms to be cardiovascular (this would include transient ischemic attacks, strokes). Pre-existing symptoms or signs such as a long-standing murmur or dyspnoea, which has been evaluated and have not changed would score as a "yes" response. Therefore, a "yes response" (affirmative response) to question 1, means the patient does not have any new cardiovascular sign or symptom, or in those with pre-existing cardiovascular illness, there has not been a worsening of their clinical status.

Q2: Is this a routine surveillance scan? This captures tests being considered for a "periodic" evaluation since a certain period of time has elapsed. The test is not being ordered due to the anticipation of changing clinical decision making or guiding therapy.

Q3: Has there been a previous TTE within the last year?

Q4: Is the test requested for suspected endocarditis with no positive blood cultures or new murmur?

The PSQ was applied to 501 studies in the internal validation cohort at a tertiary referral hospital and to 881 TTE requests at a regional hospital (external validation cohort). Two requests within the internal validation cohort and one request in the external validation cohort were not classifiable as it was not possible to collect information to answer the questionnaire. The final analysis was made in 499 TTE requests in the internal validation and 880 TTE requests in the external validation groups respectively.

Table 8- 2 shows study characteristics and appropriateness classification by group. The internal and external validation cohorts are well matched. However, the former group had a lower proportion of outpatients.

Table 8- 2 Study characteristics and appropriateness classification according to groups

	PSQ Derivation cohort (n=814)	**Internal validation cohort (n=499)**	**External validation cohort (n=880)**	**p value**
Age years (Median [IQR])	65.00 [52.00, 76.00]	67 [55, 76]	70 [58, 80]	<0.01
Male, n (%)	444 (54.6)	256 (51.3)	449 (51.0)	0.96
Outpatient, n (%)	573 (70.4)	289 (57.9)	661 (75.1)	<0.01
Referred by cardiologists, n (%)	348 (42.8)	253 (50.7)	686 (78.0)	<0.01
Appropriateness score				
Appropriate, n (%)	707 (86.9)	431 (86.4)	774 (88.0)	<0.01
RA appropriate, n (%)	68 (8.4)	37 (7.4)	75 (8.5)	
May be appropriate, n (%)	13 (1.6)	18 (3.6)	7 (0.8)	
Unclassifiable, n (%)	26 (3.2)	13 (2.6)	24 (2.7)	

The 10 most common RA indications in the internal validation group are described in Table 8-3. Inter-rater agreement for scoring between both reviewers was high (kappa=89%).

Table 8- 3 Ten most common rarely appropriate indications in prospective internal validation cohort

Indication		**#**	**Proportion**
88	Routine surveillance (<1 y) of known cardiomyopathy without a change in clinical status or cardiac exam	7	0.19
10	Initial evaluation of ventricular function (e.g., screening) with no symptoms or signs of cardiovascular disease	5	0.14
53	Transient fever without evidence of bacteraemia or a new murmur	4	0.11
60	Routine surveillance of known small pericardial effusion with no change in clinical status	3	0.08
35	Initial evaluation when there are no other symptoms or signs of valvular or structural heart disease	2	0.05
36	Re-evaluation in a patient without valvular disease on prior echocardiogram and no change in clinical status or cardiac exam	2	0.05
40	Routine surveillance (<1 y) of moderate or severe valvular stenosis without a change in clinical status or cardiac exam	2	0.05
48	Routine surveillance (<3 y after valve implantation) of prosthetic valve if no known or suspected valve dysfunction	2	0.05
74	Routine surveillance (<1 y) of HF (systolic or diastolic) when there is no change in clinical status or cardiac exam	2	0.05
79	Routine surveillance (<1 y) of implanted device without a change in clinical status or cardiac exam	2	0.05

When question 1 (no change in symptoms/no change in clinical status/no change in cardiac exam) was answered affirmatively, it had a higher OR, greater positive predictive value and positive likelihood ratio to pick up possible RA tests when compared independently to other questions. This was driven by examination of asymptomatic or stable patients (Appendix Table 6).

In the internal validation cohort, 18% of the tests had ≥2 affirmative answers (two "yes" responses). A PSQ with ≥ 2 affirmative responses had an Odds Ratio (OR) 33.96 [13.61, 84.78], sensitivity 0.84 [0.68, 0.94] and specificity of 0.87 [0.83, 0.90] for inappropriate requests.

In the external validation group (n=880), ≥2 affirmative answers were 11% of total requests; it had a sensitivity of 80%, specificity of 95% and OR 83.01 (Table 8- 4).

Table 8- 4 Diagnostic tests of the Point of Service Questionnaire for RA requests

	Sensitivity Estimate [lower-upper]	**Specificity Estimate [lower-upper]**	**Odds ratio Estimate [lower-upper]**	**Positive predictive value Estimate [lower-upper]**	**Negative predictive value**	**Positive likelihood ratio Estimate [lower-upper]**	**Negative likelihood ratio Estimate [lower-upper]**
Prospective cohort	0.84[0.68,0.94]	0.87[0.83,0.90]	33.96[13.61,84.78]	0.34[0.24,0.44]	0.98[0.97,1.00]	6.34[4.83,8.34]	0.19[0.09,0.39]
External retrospective validation cohort	0.80[0.69,0.88]	0.95[0.84,0.97]	83.3[43.13, 159.83]	0.62[0.51,0.72]	0.98[0.97,0.99]	17.41[12.46,24.32]	0.21[0.13,0.33]

The PSQ with ≥2 affirmative responses identified 84%, and 80% of the inappropriate tests in the internal and external validation cohorts respectively.

Around 20% of the TTE requests provided inadequate information and it was necessary to check the digital medical records (Table 8- 5).

Table 8- 5 Differences in time when medical record (DMR) needed to be checked.

	Internal validation cohort		**External validation cohort**	
	yes	no	yes	no
Needed DMR	95 (0.19)	404 (0.81)	168 (0.19)	712 (0.81)
Seconds (median [IQR])	120 [62, 120]	18 [14, 21]	80 [63, 98]	12 [9, 20]

Discussion:

The approach proposed in this study has been to encapsulate the essence of the RA tests into four binary questions which can be used rapidly to screen for these inappropriate requests. The questions which we have utilised are consistent with published literature with the choosing wisely programme identifying "Routine studies" and "no change in signs or symptoms" as unnecessary repeat testing[187]. Similarly, other authors identified that 54% of inappropriate requests had a TTE within the last year[59]. Although only 28% of the inappropriate studies in the derivation cohort had a TTE within the previous year, we included this question to improve the strength of our model. Finally given the high prevalence of inappropriate endocarditis requests (26% in our derivation cohort), it was included as the final question in our model.

The primary finding of this study is that the application of the PSQ is feasible, and it identifies a high proportion of inappropriate tests without the need to review all the TTE requests against the AC. An affirmative response to any of the questions increases the likelihood of a test being deemed "Rarely Appropriate", and when two (or more) of the questions were answered affirmatively, the chance of determining a test as inappropriate increased more than 33 times. The results were consistent in the different cohorts and scenarios (inpatients/outpatients, test

referred by cardiologists/non-cardiologists). We propose a PSQ based method for screening appropriateness (Figure 8- 2).

Using this model less than 1/5 of TTE requests would need to be audited against the AUC thereby minimising interruption of work flow. Nevertheless, occasional requests such as asymptomatic severe mitral regurgitation surveillance within 1 years (AC indication 45 Uncertain) or repeat echocardiography for a heart failure patient on optimal medical therapy without a change in signs on symptoms to guide therapy (AUC indication 73: Appropriate) could still be performed despite 2 affirmative responses.

The use of radiology benefit management companies (RBM) is still an important pole in the process of performance of cardiac imaging although one of the aims of the AC was to reduce the need of those companies.[189] Prior authorization and claim denials continue to be the top challenges of the process.[190] The results of this study show that the use of the questionnaire provides a transparent solution, which can be implemented with minimal delay at point-of-service, thereby minimising the need for RBMs or other middlemen.

Several attempts have been made to improve appropriateness at the point-of-order and care, implementing software to control the request of inappropriate tests.[45, 126, 184] However, the use of those tools at the point-of-order are susceptible to indication drift: for example, the real indication for testing may be RA but inactive problems are appropriate. Perhaps for this reason, the AC literature has shown little or no improvement in requesting behaviour.[126, 184].

Our proposed method could be used to facilitate appropriate use audits (simplifying to 4 questions from >200 AUC, which is useful where these data are not available in electronic format), or added to the current appropriate use process at the point-of-service. At that level, the proposed approach provides a simple screening tool to flag possibly inappropriate tests at the time of scheduling.

In our study, the prevalence of RA tests varied between 7 and 9%. These values fall on the lower end of the prevalence distribution of RA tests documented in various studies.[6, 35, 42, 45, 49, 59, 120]

A comprehensive plan detailing the management of inappropriate requests found at the point of service is lacking. At the very least, a clear and effective communication strategy needs to be in place to inform a discussion before a "rarely appropriate" test is scheduled.

This questionnaire should act as a prompt to refer to the AC rather than an absolute assessment of appropriateness. Education has resulted in increased awareness of the AC without a significant change in clinical practice.[6] It is intuitive to entrust point-of-care policing regarding the appropriateness of an investigation to those with the greatest experience. However, such a policy may result in interruption of workflow, delays and a greater burden on already busy echocardiography units.

Our results show that the use of 4 simple binary questions identifies RA tests with a high sensitivity and specificity. The questionnaire identifies potential RA requests which can then be confirmed by reference to the AC (Available online, in-print or as a mobile phone applications) or by discussing the case with a Cardiologist at the echocardiographic laboratory or with the referring physician.

The use of the questionnaire may face some challenges. Firstly, the need to review medical records to adjudicate appropriateness when inadequate information was provided on request forms is a potential limitation. However, corroboration of clinical history is a common clinical practice in all imaging units as it is mandatory to establish the question being asked of an investigation and is essential to implementing a Bayesian approach to reporting. This raises a second issue, which is quality control of echocardiography requests. We identified 20% of requests as inadequate requiring further corroboration of clinical history. This result was similar to recently published data, which found in a review of 1303 requests that 26.2% were inadequate to determine adherence to AC.[191] The study concluded that the top three reasons for inadequacy was failure to report change in clinical status or cardiac exam, date of prior echocardiogram and type and severity of valvular lesion. Though the last 2 would be easily accessible in any echocardiographic laboratory (assuming the previous studies were performed within that laboratory), the first failure is critical in determining appropriateness. Clearly, it follows that access to electronic medical records is a necessary and essential component of an echocardiographic laboratory's workflow.

Secondly, the possible difference in referral patterns between hospitals and private labs, may impact on the positive predictive value of our questionnaire. The Private/ Public divide varies tremendously across countries and to date has not been assessed in regards to compliance with AC. Though our criteria address the issue of awareness and simplicity perhaps, the greatest challenge facing an overburdened medical infrastructure is the systemic dependence of investigations.

Furthermore, although our rate of inappropriate use is lower than reported in other institutions, and could be seen as a limitation, we overcame this issue by performing an analysis of over 800 requests in the derivation process. We also validated this questionnaire in over 1200 patients (internal and external validation cohorts).

Finally, previous researchers have sought to differentiate the appropriateness of a study from its clinical utility arguing a RA test does not necessarily mean a clinically useless one nor does an appropriate request always correspond with a useful one.[35, 59] Parker Ward, demonstrated that 17% of inappropriate tests had "new important TTE abnormalities" and Matulevicius, showed that 21.7 % of Rarely Appropriate tests led to an active change in management. Thus, identification of inappropriate tests is a step on the path to improving quality and appropriateness in cardiovascular imaging; decision-making has to be informed by individual characteristics.

How we handle inappropriate requests will ultimately have financial and clinical implications. The proportion of inappropriateness is between 7% and 23%.[6, 187] Our study shows an inappropriate rate of 7.4% to 8.5%. Experience from elsewhere in Australia has demonstrated the inappropriate rate of 20%.[192] In 2015, the total Medicare reimbursement for transthoracic echocardiography was approximately AUD$186.0 million.[76] Assuming a rate of inappropriate echocardiography between 7% and 20%, the cost to the Australian health system of inappropriate transthoracic echocardiography would be between AUD$13.0 - 37.2 million. The healthcare costs are clearly proportionate to the use of transthoracic echocardiography and prevalence of appropriate use.

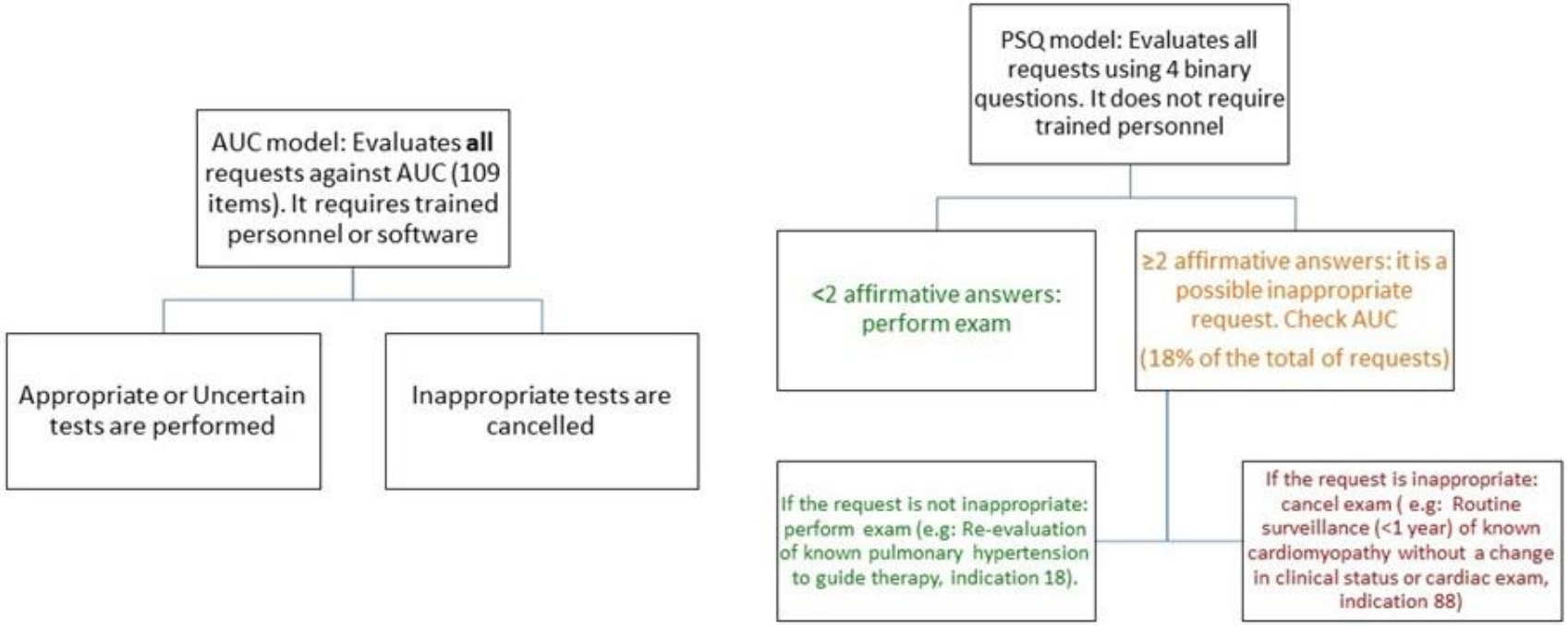
A mandatory AC score (Appropriate or otherwise) or point-of-service score tied to funding would enhance compliance with AC and enable continuous auditing of resource utilisation. There are > 100 categories of appropriateness and the incremental workflow issues are prohibitive. We proposed an alternative approach where response to the PSQ serves as a less cumbersome beacon of appropriateness.

Conclusion

We have demonstrated that $\geq$2 affirmative answers at a simple PSQ detect a high proportion of RA tests. This approach can be used as a red flag for inappropriate examinations and a prompt to further discussion about the suitability for testing in individual patients. We propose this PSQ can be a quality control tool that captures the majority of inappropriate use, in the absence

of the infrastructure that supports AC in North America, and a simple marker for departmental and regional audits.

Figure 8- 2 Comparison between AC model and PSQ model.



Chapter 9. Usefulness of Hand-held echocardiography in inappropriate tests

The research contained within this chapter has been published as[9]:

- Pathan F, Fonseca R and Marwick TH. Usefulness of Hand-Held Ultrasonography as a Gatekeeper to Standard Echocardiography for "Rarely Appropriate" Echocardiography Requests. *Am J Cardiol*. 2016;118:1588-1592.*

  **Drs Pathan and Fonseca should be regarded as joint first authors.*

Preface

The previous chapter presented a strategy to identify inappropriate requests at the point of service using an easy questionnaire, which can be responded to with no impact on the workflow of the site. The point-of-service questionnaire is a "red flag" to inappropriate tests, which, instead of being cancelled, should be analysed in a deeper way to determine the need of the test. However, the needs of clinicians leading to test orders, as exemplified in Chapter 6, in chapter 7, this chapter aims to determine if the use of hand-held echocardiography could help clinicians by providing the information they need to make informed decisions.At the same time, the approach could avoid the performance of inappropriate tests at the echo laboratory, impacting the workload of the service and expenditure.

Abstract

Adoption of Appropriate Use Criteria (AUC) has not had a major impact on the frequency of "rarely appropriate" tests, with the rarely appropriate tests rate remaining at ~20% in most institutions. We sought whether access to Hand-Held Ultrasound (HHU) could be an alternative means of reducing "rarely appropriate" requests. We compared two approaches to rarely appropriate requests; "standard echocardiography" (SE) as requested (control) and Hand-Held Ultrasound as a gatekeeper (HHU).

Methods and results

Patients were followed up for 6 months and assessed for endpoints including time until scan, repeat echocardiography/ cost of either strategy, new major pathology and change in management. The most common rarely appropriate requests in both groups were assessment of infective endocarditis without positive blood cultures and precordial murmur evaluation in absence of any other symptoms or signs of cardiovascular disease.

The groups had comparable age, gender, requesting physician, and inpatient vs. outpatient distribution. HHU led to a 59% reduction in rarely appropriate requests requiring SE. HHU significantly reduced time to decision for inpatients (0 [IQR: 0,1] vs 2 days [IQR: 1,4], $p<0.001$) and total cost of echocardiography ($109±86 vs $181±37, $p<0.001$). New major pathology was identified in 29% and 23% of HHU and SE respectively.

There was no difference with respect to change in management.

Conclusion.

HHU can be an effective gatekeeper to SE for rarely appropriate echocardiograms, reducing time to echocardiography and cost while satisfying the referring physician and avoiding repeat requests for SE. HHU provides a safety net, which identifies potential important findings in rarely appropriate requests.

Keywords: Handheld ultrasonography, Appropriate Use

Introduction

The availability and safety profile of echocardiography have contributed to overutilization of this technique, with echocardiography being a significant contributor to the growth of Medicare reimbursements to cardiologists from USD 1.6 to 5.4 billion between 2000 and 2006 [18]. The Appropriate Use Criteria (AUC) and educational initiatives have been components of efforts to restrain inappropriate use and costly care [26, 193].

Despite a decrease in reimbursements by 33% for echocardiography between 2006 and 2010[173], there is limited evidence to suggest that there is a decline in the rate of inappropriate echocardiography [6]. Indeed, interventions targeting education and the use of point of order applications have been applied with varying degrees of success [120, 184].

A problem with the AUC process is that restriction of tests may be contrary to the guidelines[5] and risks missing significant pathology. Parker Ward demonstrated that 17% of inappropriate studies have "*new important transthoracic echocardiography (TTE) abnormalities*" [59]; 22% of rarely appropriate tests resulted in an active change in management[35].

In the context of these observations, it is hard to enforce an overarching prohibition on rarely appropriate requests. Although previously unconnected, hand-held ultrasonography (HHU) has developed over a similar timeline to the AUC. Validation studies against standard TTE (SE), show high levels of agreement for morphology, functional and valvular assessment (kappa 0.90 to 0.99) [194]. Comparisons of HHU to physical examination have demonstrated the superiority of the former and its cost-effectiveness [195].

We hypothesised that HHU could be applied as a gatekeeper to SE for rarely appropriate examinations and that such a strategy would: Reduce numbers of SE performed and cost, identify relevant pathology which would have been missed if the rarely appropriate tests were cancelled and facilitate decision making and patient management.

Methods

This was a case-control study designed to compare a Hand-Held Ultrasonography based approach to rarely appropriate tests to the current SE based system. It was performed across two hospitals and included both inpatients and outpatients. Rarely appropriate requests were identified using the AUC[26]. If the clinical history was inadequate to ascertain appropriateness, medical records were consulted. The process of cancelling "rarely appropriate" requests is not

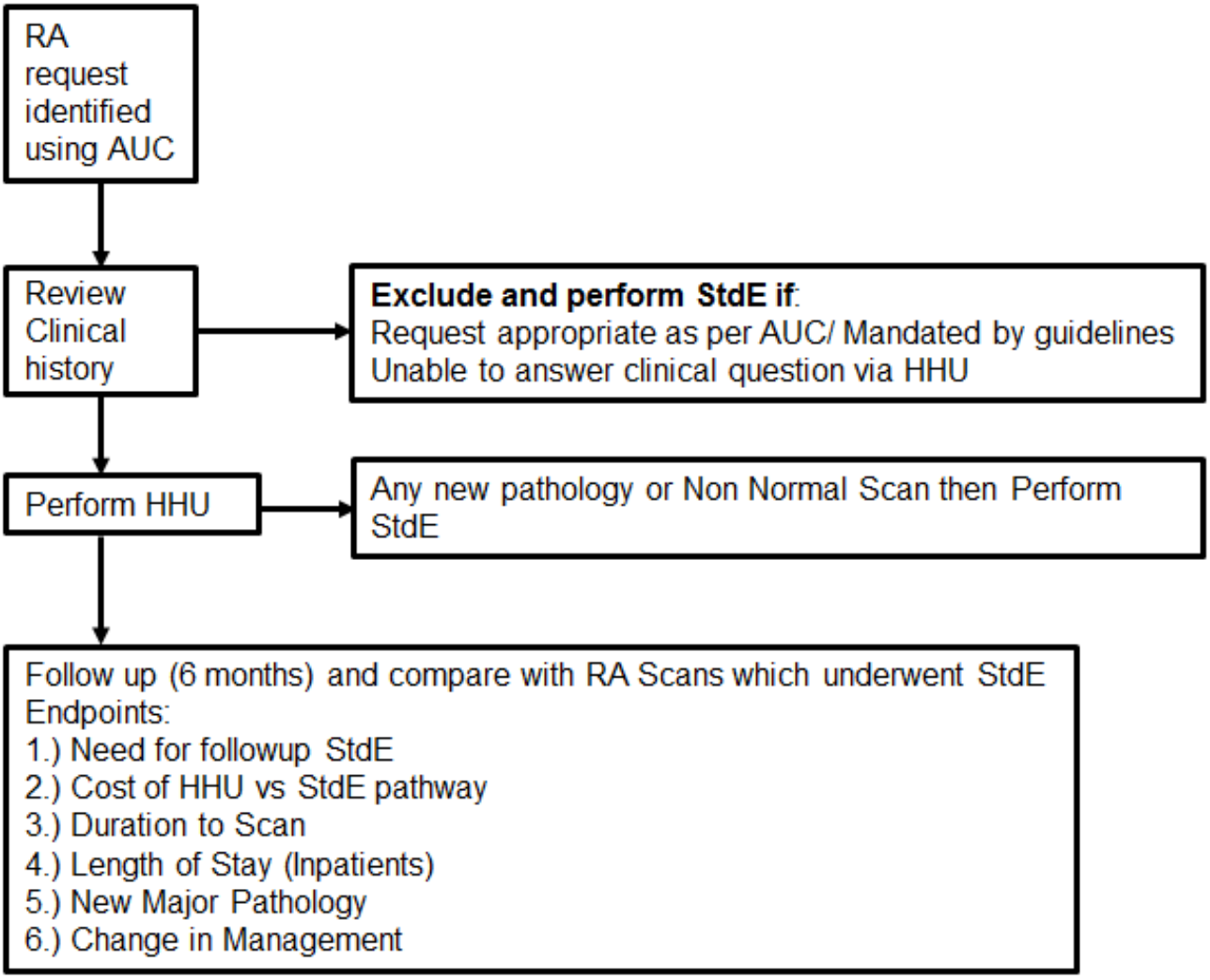
systematic and is dependent on inpatient and outpatient waiting lists. We performed HHU on requests, which were destined to be cancelled.

The study was performed between March 2015 and December 2015. All rarely appropriate requests during this period were assessed for eligibility. Requests were excluded from analysis if deemed "appropriate" following review of medical records or if tests were mandated by guidelines. Requests beyond the scope of HHU were also excluded. The resulting cohort of HHU cases was case matched 2:1 to a cohort of rarely appropriate requests which had undergone SE between 2013 and 2015.

Eligible patients with rarely appropriate requests received a cardiology consultation and HHU examination performed by a cardiologist (Figure 9- 1).

If the consult suggested the request was appropriate, then a SE was conducted and these patients were excluded. The study was approved by the Tasmanian Human Research Ethics Committee.

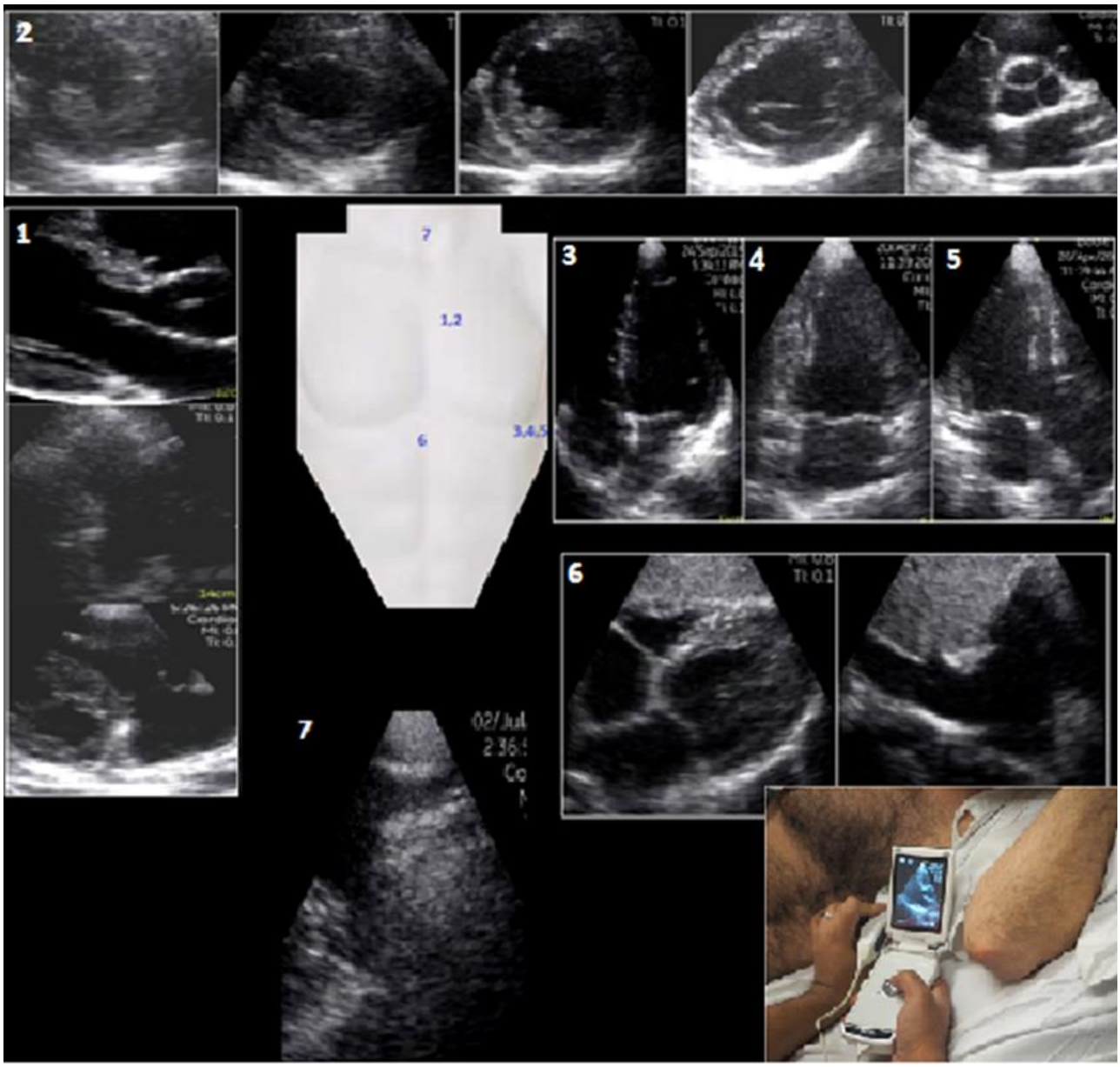
Figure 9- 1 Study design



AUC= Appropriate Use Criteria, RA= Rarely Appropriate, SE= Standard Echocardiographic examination

HHU was performed using the GE Vscan V1.2 handheld device (GE Health Care, Milwaukee, WI). The HHU study protocol involved 2D grey scale and colour Doppler images across all standardised echocardiographic (Figure 9- 2). Linear or area measurements were made as appropriate. Spectral Doppler data was not obtained as this characteristic is not available on the HHU. The severity of disease was approximated using 2D signs of severity (e.g. leaflet excursion, chamber dilatation), size, and duration of the colour jet / proximal convergence zone.

Figure 9- 2 Hand-held ultrasound protocol



1,2) Parasternal long and short axis views including short axis sweep for left ventricular function and regional wall motion assessment.

3,4,5) Apical views as illustrated and 5 chamber for aortic valve assessment.

6) Subcostal- long axis and inferior vena cava view

7) Aortic Arch view. Colour Doppler applied to all views.

All HHU examinations were limited to < 10 minutes' duration. Results of the HHU study were communicated to the treating team, documented in clinical notes or provided as a short report for outpatients.

Eligible patients went on to have a formal SE if the HHU suggested a full study was warranted: Non-diagnostic HHU, Any HHU abnormality in patients without a previous echocardiogram and any new changes in patients with a previous examination.

The HHU cohort was case-matched to a cohort of patients with rarely appropriate tests, who underwent the reference SE. The SE examination was performed on 3 machines: GE Vivid 9 (GE Health Care, Milwaukee, WI), Phillips IE33 (Philips Medical Systems, Andover, Massachusetts) and Acuson SC2000 (Siemens, North Rhine-Westphalia, Germany). All SE were performed over 45- 60 minutes by an experienced sonographer, and interpreted by a cardiologist.

Patients in both HHU and SE arms were followed up for 6 months and evaluated for the endpoints of: repeat TTE, the cost of care, time to scan, the length of stay (inpatients), new major TTE abnormality/ incidental findings and change in management. The cost of additional scans was incorporated into each arm.

The cost of SE was standardised to $230 pre-examination as per the Australian Medicare benefits schedule[91]. The cost of the hand-held device is $6000 USD - assuming a depreciation of $750/ year and its use in 100 rarely appropriate requests per year, the cost per scan of the device would be $7.50 per rarely appropriate scan. We added to this the cost of storage of images, 15 minutes' time for the cardiologist to perform consultation, echocardiography and generate a report to arrive at an estimated cost of $38 USD per scan. This approach is consistent with previous attempts to itemise a cost for HHU accounting for geographical variations (Table 9- 1)[194-202].

Table 9- 1 Cost benefit studies of hand held ultrasonography vs standard echocardiography. a

Author	**Year**	**Country**	**Device**	**HHU Cost**	**Operator Cost**	**Total HHU Cost (Device+Operator)**	**Total StdE Cost**
Vourvouri	2003	Netherlands	SonoHeart	$5	$116[b]	$121	$213
Greaves	2005	UK	OptiGo	NR	NR	$8	$128
Galasko	2006	UK	OptiGo	NR	NR	$65- $50	$258
Trambaiolo	2007	Italy	OptiGo	$14	$26	$40	$122
Gianstefani	2013	UK	VScan	$1	$33[c]	$34	$137
Kitada	2013	Japan	VScan	$2	NR	$2	$88
Khan	2014	USA	VScan	NR	NR	NR	$800
Mehta	2014	USA	VScan	$8	$11	$19	$1511[d]
Kini	2015	USA	VScan	$9	$0- 23	$9- 32	$73
Pathan	**2016**	**Australia**	**VScan**	**$8**	**$30**	**$38**	**$173**

NR= Not Reported

a.) All costs converted to US dollars based on conversion rate at time of publication, b.) Including Cardiology Consult, c.) £25 + 8 (Operator + Hospital costs), d.) $162 (Professional charge) + $1349 (Facility charge)

Major pathology was defined as [59]: moderate or greater (left ventricular dysfunction, valvular regurgitation/ stenosis, pulmonary hypertension, diastolic dysfunction); a regional wall motion abnormality; right ventricular dysfunction and moderate or greater pericardial effusion, thrombus or vegetation.

Changes in management were characterised as an active change in care including medication changes, subspecialty consultation, surgery or invasive procedures, diagnostic testing, change in the level of care, cancellation of initially planned procedure or intervention.

Continuation of care: no escalation or de-escalation of current care, following direct communication about TTE results to patients and documentation by providers in the medical record.

No change in care: no change in therapy or documentation of reassurance about TTE findings after TTE was performed, or the next step in management was already documented and the plan in place in place before the TTE result or results were not accessed, acknowledged or noted in further correspondence or discharge summaries.

The statistical analysis was performed using R version 3.2.2 software.[79] Baseline characteristics and outcomes were compared for the HHU and SE groups. Categorical variables and outcomes were compared using a chi-square test and continuous variables using the Student's t-test with a $p < 0.05$ considered statistically significant.

Results

A total of 872 Echocardiography request forms were audited for appropriateness. Based on the AUC, 93 (10.6%) requests were deemed inappropriate. Routine follow-up of pulmonary hypertension, which is mandated by guidelines, was responsible for 37 of these requests, which were then excluded. Of the remaining 56 requests, 15 were excluded due to: Inability to answer clinical question with HHU (3 requests), appropriate following cardiology consultation (2) and unable to attend an appointment for echocardiography (10). The 41 remaining cases included in the analysis were case matched 2:1 with a retrospective cohort.

A total of 123 (41 HHU, 82 SE) patients in whom rarely appropriate tests were requested, were included in this study. There were no significant differences in the clinical or the request characteristics of the two groups (Table 9- 2).

A cardiologist was the requesting physician in 24% of the HHU and 32% of SE examinations (p=0.53). The remaining requests in each arm were ordered by a combination of medical and surgical physicians.

"Rarely appropriate" studies for infective endocarditis, routine follow-up for heart failure and evaluation of a precordial murmur or ventricular function in the absence of cardiovascular signs and symptoms were the most common requests in both arms.

Table 9- 2 Patient and Request Characteristics at Baseline

Patient Characteristics	**HHU (n= 41)**	**SE (n= 82)**	**p**
Age (median [IQR]) (years)	62 [52, 73]	61 [48, 72]	0.60
Male	23 (56%)	40 (49%)	0.57
Hypertension	21 (51%)	38 (46%)	0.75
Diabetes mellitus	6 (15%)	11 (13%)	1.00
Valvular heart disease	6 (15%)	15 (18%)	0.80
Atrial fibrillation	4 (10%)	12 (15%)	0.58
Ischemic heart disease	12 (29%)	15 (18%)	0.25
Heart failure	7 (17%)	12 (15%)	0.93
Lung disease	4 (10%)	14 (17%)	0.42
Renal disease	4 (10%)	8 (10%)	1.00
Liver disease	2 (5%)	10 (12%)	0.33
Systemic disease	4 (10%)	8 (10%)	1.00
Requested by Cardiologist	10 (24%)	26 (32%)	0.53
Setting Ambulatory care/Outpatient	20 (49%)	48 (58%)	0.41

IQR= interquartile range, HHU= Hand held Ultrasound (Echocardiography), SE= Standard transthoracic echocardiography

Table 9- 3 outlines the result of all endpoints. The time to scan and subsequent decision for inpatients was significantly shorter in the HHU arm [0 vs. 2 days ($p<0.001$)]. There was no significant difference in relation to duration of inpatient stay between HHU 14 days and SE 9 days ($p=0.35$). There was no significant difference between the time to scan for outpatients between the HHU and SE [32 vs. 35 ($p= 0.54$)].

As expected, more patients in the HHU arm, 41% required follow-up SE study, which was mandated due to results of HHU and our safety protocol [new major pathology (12 cases), minor pathology (3) and non-diagnostic HHU (2)]. In the SE arm, 11% required follow-up echocardiography. However, 5 of these were due to new clinically appropriate indications over the course of 6 months' follow-up and these costs were not included in the SE costing assessment. The mean cost of the HHU strategy was significantly less than the SE arm [$109 vs. 181 USD ($p<0.001$)], resulting in a saving of $72 per study.

During 6 months follow-up of the 123 cases (HHU and SE) 31/123 (25%) of all rarely appropriate requests had new major pathology. There was no significant difference ($p=0.15$) between HHU arm and the SE arm with respect to major pathology. Only 1 patient in the HHU arm and 0 patients in the SE arm had a documented incidental finding - a liver mass requiring an abdominal ultrasound.

Of 123 requests, which were deemed "rarely appropriate", 15% of results led to an active change in management using either HHU or SE. Furthermore, 64% of results led to the continuation of care. Thus only 21% of investigations resulted in absolutely no change in management. There were no significant differences between the two groups with regards to change in management ($p=0.27$).

Table 9- 3 Endpoints

Endpoints	**HHU (n= 41)**	**SE (n= 82)**	**p**
Time to scan days Inpatients, (median [IQR])	0 [0, 1]	2 [1, 4]	**<0.001**
LOS days Inpatients, (median [IQR])	14 [7, 28]	9 [5, 21]	0.35
Time to scan days Outpatients, (median [IQR])	31 [31, 32]	35 [21, 108]	0.54
Follow up SE within 6 months n (%)	17 (41%)	9 (11%)	**<0.001**
Average cost USD, (mean (sd))	109 (86)	181 (37)	**<0.001**
New Major Pathology	12 (29%)	19 (23%)	0.15
Change in management			0.27
No Change	8 (19%)	18 (22%)	
Continuation of Care	24 (58%)	55 (67%)	
Active change in management	9 (22%)	9 (11%)	0.18
Change in Management or Continuation of Care	33 (80%)	64 (78%)	0.94

USD= US Dollar (Costs were calculated in Australian dollars and converted to US dollars. 1 Australian dollar=0.75 USD), HHU= Hand Held Ultrasound, IQR= interquartile range, LOS= length of stay, sd= Standard Deviation, SE= Standard transthoracic echocardiography

Discussion

The results of this study show that despite being classified as rarely appropriate, 25% of total requests had new major TTE abnormalities, which would have been missed if these tests were cancelled. The use of HHU to screen rarely appropriate echocardiography requests led to a 59% reduction in SE studies, without any compromise in outcomes related to management. To our knowledge this is the first study, which specifically provides an alternative imaging strategy to the performance or cancellation of rarely appropriate tests.

The current rate of "rarely appropriate" echocardiography remains at 10-20% [6, 203]. Various strategies have been proposed to enforce the AUC, including education, point of order software, and radiology benefit management [99]. Ongoing rarely appropriate requests reflect the concern that these echocardiograms may still be clinically useful and influence management [35, 59, 204]A HHU based approach would enable application of the AUC and reduce the performance of inappropriate echocardiograms while simultaneously creating a safety net to identify significant findings. Indeed, a HHU-backed approach to the AUC was a cost-saving strategy, with a saving of $72 per study.

All inpatient referrals were non-cardiology admissions, and the longer duration of stay was unexpectedly higher than the expected 4 day length of stay [205]. The use of a HHU strategy in inpatients did not influence length of stay, reflecting the complex patient phenotype where that duration of stay is related to clinical factors.

Although not specifically devoted to use of HHU in rarely appropriate tests, previous work supports the use of HHU in common situations associated with rarely appropriate tests. In a study of HHU and repeat TTE in 105 adult patients undergoing follow-up echocardiography, HHU showed good to excellent correlation with TTE[202]. In addition, a HHU protocol could save between $41-64 per study. Twelve percent of these patients were deemed rarely appropriate, but a subgroup analysis was not performed.

Our study was not a direct comparison of HHU vs. SE, and cannot address the question of echocardiographic false negative scans, although the question of validation has been explored in great detail, including in recent studies [194, 202]. A consensus statement published by the American Society of Echocardiography found that LV enlargement, LV hypertrophy, LV systolic function, LA enlargement, RV enlargement, RV systolic function, pericardial effusion and IVC size have all been accurately detected [206].

This is a case-control study; despite the patients in both arms being well matched, the lack of a randomised controlled trial may have resulted in potential selection bias. A randomised study should be considered, with head-to-head comparison of HHU with SE. From a practical standpoint, a cardiologist may not always be able to provide a consultation. Finally, the categorization of "change in management" is a difficult task. In using the criteria proposed by Matulevicius et al. we recognise the limitations of relying on medical documentation and correspondence to differentiate between "No change in Management" and "Continuation of Care".

Conclusion

Hand-held ultrasonography provides a safety net and cost-saving strategy that identifies potential important findings in inappropriate requests.

Chapter 10. Summary and conclusion

The studies contained in this PhD thesis help us to comprehend that in Australia, the cardiac imaging techniques, particularly echocardiography, are not used "appropriately". It is estimated that between 7% and 23% of transthoracic echocardiography requests are performed for "inappropriate" reasons (this estimation is only for transthoracic echocardiography, not other echocardiography techniques, cardiac tomography, cardiac resonance, or SPECT)[6]. In 2015, Medicare reimbursed (only for transthoracic echocardiography) ~AUD$186.0 million[8, 76]. Therefore, the cost to the Australian health system of inappropriate transthoracic echocardiography would be between AUD$13.0 and 37.2 million[8]. Consequently, it is mandatory to improve the appropriate use of cardiac imaging.

Medical associations in the USA developed the Appropriate Use Criteria to help clinicians to improve appropriate use of imaging by choosing "appropriate" clinical scenarios[26]. Part of the work as this investigation was to assess if the Appropriate Use Criteria utilised by the American Associations were suitable to use in Australia. However, the studies included in this thesis indicated that the Appropriate Use Criteria would face several challenges, if used as currently designed, in Australia. These problems are based on the following facts found during in this research:

First, the Criteria are not concordant with the cardiac guidelines, which can lead to misunderstandings and misinterpretations among our clinicians[5]. Second, the Criteria have not changed the physicians' requesting behaviour over time because we did not find a decrease in the proportion of inappropriate tests since 2005[6]. The third factor is that we did not find differences in survival time/readmission of patients with Heart Failure if doctors request echocardiograms following the criteria[7], and lastly, the results of the qualitative research suggest there is a mismatch between the clinical reasoning of physicians and the AUC for echocardiography.

However, it remains necessary to control the use of cardiac imaging and to decrease the amount of "inappropriate" testing. Because we see that the Appropriate Use Criteria would not help us to achieve the goal of "appropriate" use, we designed a questionnaire (4 binary (yes/no) questions) for utilisation in the echo laboratory, which helps us to determine a high proportion of possible "inappropriate" requests. Having at least two affirmative answers to the

questionnaire gives us a high probability of having an "inappropriate" echo request[8]. Instead of performing the expensive, time-consuming echocardiogram (which costs around AUD$250 and needs around 45 minutes), we found that we could perform a hand-held echocardiogram to respond to the "inappropriate" requests[9]. This approach can be an effective gatekeeper to the standard echocardiogram for "inappropriate" tests, reducing time to echocardiography and cost while satisfying the referring physician and avoiding repeat requests for standard echocardiography[9].

The "appropriate" use of echocardiography is perhaps a hard aim to achieve. However, its achievement will be the result of the joint effort of clinicians, policy makers, government, and patients. Although this thesis does not give us all the possible solutions to this problem, it gives us an important tool to use at to decrease the overuse of echocardiography, answering the questions that our physicians need to manage their patients.

Bibliography

1. Fonseca R and Marwick TH. Appropriateness and outcomes: is it time to adopt appropriate use criteria outside of North America? *Heart*. 2014;100:357-358.

2. Fonseca R and Marwick TH. How I do it: judging appropriateness for TTE and TEE. *Cardiovasc Ultrasoun*. 2014;12:22.

3. Centers for M and Medicaid Services HHS. Medicare program; hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system and fiscal year 2015 rates; quality reporting requirements for specific providers; reasonable compensation equivalents for physician services in excluded hospitals and certain teaching hospitals; provider administrative appeals and judicial review; enforcement provisions for organ transplant centers; and electronic health record (EHR) incentive program. Final rule. *Federal register*. 2014;79:49853-50536.

4. Fonseca R, Otahal P, Wiggins N and Marwick TH. Growth and geographical variation in the use of cardiac imaging in Australia. *Intern Med J*. 2015;45:1115-27.

5. Fonseca R, Negishi K and Marwick TH. What is the evidence status of Appropriate Use Criteria (AUC)? Insight from a matching exercise with the guidelines for echocardiography. *Intern Med J*. 2015;45:864-9.

6. Fonseca R, Negishi K, Otahal P and Marwick TH. Temporal changes in appropriateness of cardiac imaging. *J Am Coll Cardiol*. 2015;65:763-73.

7. Fonseca R, Otahal P, Galligan J, Neilson S, Huynh Q, Saito M, Negishi K and Marwick TH. Association of survival time with transthoracic echocardiography in stable patients with heart failure: Is routine follow-up ever appropriate? *Int J Cardiol*. 2017;230:619-624.

8. Fonseca R, Pathan F and Marwick TH. Development and validation of a screening tool for the identification of inappropriate transthoracic echocardiograms. *BMJ Open*. 2016;6:e012702.

9. Pathan F, Fonseca R and Marwick TH. Usefulness of Hand-Held Ultrasonography as a Gatekeeper to Standard Echocardiography for "Rarely Appropriate" Echocardiography Requests. *Am J Cardiol*. 2016;118:1588-1592.

10. Jakovljevic M, Potapchik E, Popovich L, Barik D and Getzen TE. Evolving Health Expenditure Landscape of the BRICS Nations and Projections to 2025. *Health Econ*. 2016:n/a-n/a.

11. OECD. *Fiscal Sustainability of Health Systems*: OECD Publishing; 2015.

12. Brownlee S, Chalkidou K, Doust J, Elshaug AG, Glasziou P, Heath I, Nagpal S, Saini V, Srivastava D, Chalmers K and Korenstein D. Evidence for overuse of medical services around the world. *Lancet*. 2017.

13. Bennetts R, Dernie E, Gray B, Majchrzak-Smith A, Miller J, Xu D, Webster A and Welfare AIoHa. 25 years of health expenditure in Australia, 1989-90 to 2013-14. 2016.

14. Baltagi BH, Lagravinese R, Moscone F and Tosetti E. Health Care Expenditure and Income: A Global Perspective. *Health Econ*. 2016:n/a-n/a.

15. Arrow KJ. Uncertainty and the Welfare Economics of Medical-Care. *Am Econ Rev*. 1963;53:941-973.

16. Shaw LJ, Marwick TH, Zoghbi WA, Hundley WG, Kramer CM, Achenbach S, Dilsizian V, Kern MJ, Chandrashekhar Y and Narula J. Why all the focus on cardiac imaging? *JACC: Cardiovascular Imaging*. 2010;3:789-794.

17. Medicare Payment Advisory Commission. *Report to the Congress: Medicare Payment Policy*. Washington, D.C.: Medicare Payment Advisory Commission; 2005.

18. Shaw LJ, Marwick TH, Zoghbi WA, Hundley WG, Kramer CM, Achenbach S, Dilsizian V, Kern MJ, Chandrashekhar Y and Narula J. Why all the focus on cardiac imaging? *JACC Cardiovasc Imaging*. 2010;3:789-94.

19. Hackbarth GM, Reischauer RD and Miller ME. MedPAC recommendations on imaging services. 2006.

20. Medicare Payment Advisory Commission. *Report to the Congress: New Approaches in Medicare*. Washington, D.C.: Medicare Payment Advisory Commission; 2004.

21. Medicare Payment Advisory Commission. *Report to the Congress: Growth in the Volume of Physician Services*. Washington, DC: Medicare Payment Advisory Commission; 2004.

22. Medicare Payment Advisory Commission. *Report to the Congress: Medicare Payment Policy*. Washington, D.C.: Medicare Payment Advisory Commission; 2003.

23. Wennberg JE, Cooper MM, Birkmeyer JD, Bronner K, Bubolz TA, Campbell DE, Fisher ES, O'Connor GT, Poage JF, Sharp SM, Skinner JS, Stukel TA and Wennberg DE. The Quality of Medical Care in the United States: A Report on the Medicare Program: The Dartmouth Atlas of Health Care. 1999.

24. Wennberg DE, Birkmeyer JD, Birkmeyer NJ, Lucas FL, Malenka DJ, McGrath PD, Lurie JD, O'Connor GT, Quinton HB, Shawver TA and Siewers AE. The Dartmouth Attlas of Cardiovascular Health Care. 1999.

25. Gonzalez JA, Balfour PC, Jr. and Shaw PW. The Cardiac Imaging Specialist: Pursuing Further Training. *J Am Coll Cardiol*. 2015;66:2563-8.

26. American College of Cardiology Foundation Appropriate Use Criteria Task F, American Society of E, American Heart A, American Society of Nuclear C, Heart Failure Society of A, Heart Rhythm S, Society for Cardiovascular A, Interventions, Society of Critical Care M, Society of Cardiovascular Computed T, Society for Cardiovascular Magnetic R, Douglas PS, Garcia MJ, Haines DE, Lai WW, Manning WJ, Patel AR, Picard MH, Polk DM, Ragosta M, Ward RP and Weiner RB. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 Appropriate Use Criteria for Echocardiography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart

Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, and Society for Cardiovascular Magnetic Resonance Endorsed by the American College of Chest Physicians. *J Am Coll Cardiol*. 2011;57:1126-66.

27. Fitch K, Bernstein SJ, Aguilar MD, Burnand B and LaCalle JR. The RAND/UCLA appropriateness method user's manual. 2001.

28. Hendel RC, Patel MR, Allen JM, Min JK, Shaw LJ, Wolk MJ, Douglas PS, Kramer CM, Stainback RF, Bailey SR, Doherty JU and Brindis RG. Appropriate use of cardiovascular technology: 2013 ACCF appropriate use criteria methodology update: a report of the American College of Cardiology Foundation appropriate use criteria task force. *J Am Coll Cardiol*. 2013;61:1305-17.

29. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J and Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care*. 1986;2:53-63.

30. Patel MR, Spertus JA, Brindis RG, Hendel RC, Douglas PS, Peterson ED, Wolk MJ, Allen JM, Raskin IE and American College of Cardiology F. ACCF proposed method for evaluating the appropriateness of cardiovascular imaging. *J Am Coll Cardiol*. 2005;46:1606-13.

31. Brindis RG, Douglas PS, Hendel RC, Peterson ED, Wolk MJ, Allen JM, Patel MR, Raskin IE, Hendel RC, Bateman TM, Cerqueira MD, Gibbons RJ, Gillam LD, Gillespie JA,

Hendel RC, Iskandrian AE, Jerome SD, Krumholz HM, Messer JV, Spertus JA, Stowers SA, American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working G, American Society of Nuclear C and American Heart A. ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging (SPECT MPI): a report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group and the American Society of Nuclear Cardiology endorsed by the American Heart Association. *J Am Coll Cardiol*. 2005;46:1587-605.

32. Douglas PS, Khandheria B, Stainback RF, Weissman NJ, Brindis RG, Patel MR, Khandheria B, Alpert JS, Fitzgerald D, Heidenreich P, Martin ET, Messer JV, Miller AB, Picard MH, Raggi P, Reed KD, Rumsfeld JS, Steimle AE, Tonkovic R, Vijayaraghavan K, Weissman NJ, Yeon SB, Brindis RG, Douglas PS, Hendel RC, Patel MR, Peterson E, Wolk MJ, Allen JM, American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working G, American Society of E, American College of Emergency P, American Society of Nuclear C, Society for Cardiovascular A, Interventions, Society of Cardiovascular Computed T, Society for Cardiovascular Magnetic R, American College of Chest P and Society of Critical Care M. ACCF/ASE/ACEP/ASNC/SCAI/SCCT/SCMR 2007 appropriateness criteria for transthoracic and transesophageal echocardiography: a report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group, American Society of Echocardiography, American College of Emergency Physicians, American Society of Nuclear Cardiology, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, and the Society for

Cardiovascular Magnetic Resonance endorsed by the American College of Chest Physicians and the Society of Critical Care Medicine. *J Am Coll Cardiol*. 2007;50:187-204.

33. Douglas PS, Khandheria B, Stainback RF, Weissman NJ, Peterson ED, Hendel RC, Stainback RF, Blaivas M, Des Prez RD, Gillam LD, Golash T, Hiratzka LF, Kussmaul WG, Labovitz AJ, Lindenfeld J, Masoudi FA, Mayo PH, Porembka D, Spertus JA, Wann LS, Wiegers SE, Brindis RG, Douglas PS, Hendel RC, Patel MR, Peterson ED, Wolk MJ, Allen JM, American College of Cardiology F, American Society of E, American College of Emergency P, American Heart A, American Society of Nuclear C, Society for Cardiovascular A, Interventions, Society of Cardiovascular Computed T and Society for Cardiovascular Magnetic R. ACCF/ASE/ACEP/AHA/ASNC/SCAI/SCCT/SCMR 2008 appropriateness criteria for stress echocardiography: a report of the American College of Cardiology Foundation Appropriateness Criteria Task Force, American Society of Echocardiography, American College of Emergency Physicians, American Heart Association, American Society of Nuclear Cardiology, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, and Society for Cardiovascular Magnetic Resonance endorsed by the Heart Rhythm Society and the Society of Critical Care Medicine. *J Am Coll Cardiol*. 2008;51:1127-47.

34. Wolk MJ, Bailey SR, Doherty JU, Douglas PS, Hendel RC and Kramer CM. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 Multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease: a report of the American college of cardiology foundation appropriate use criteria task force, American heart association, American society of echocardiography, American society of nuclear cardiology, heart failure society of America, heart rhythm society, society for

cardiovascular angiography and interventions, society of cardiovascular computed tomography, society for cardiovascular magnetic resonance, and society of thoracic surgeons. *J Am Coll Cardiol*. 2014;63:380-406.

35. Matulevicius SA, Rohatgi A, Das SR, Price AL, DeLuna A and Reimold SC. Appropriate use and clinical impact of transthoracic echocardiography. *JAMA Intern Med*. 2013;173:1600-7.

36. Koshy TP, Rohatgi A, Das SR, Price AL, deLuna A, Reimold N, Willett K, Reimold SC and Matulevicius SA. The association of abnormal findings on transthoracic echocardiography with 2011 Appropriate Use Criteria and clinical impact. *Int J Cardiovasc Imaging*. 2015;31:521-8.

37. Dudzinski DM, Bhatia RS and Weiner RB. Assessing the clinical impact of appropriate echocardiograms. *JAMA Intern Med*. 2014;174:1195-6.

38. Aggarwal NR, Wuthiwaropas P, Karon BL, Miller FA, Pellikka PA and American College of Cardiology F. Application of the appropriateness criteria for echocardiography in an academic medical center. *J Am Soc Echocardiogr*. 2010;23:267-74.

39. Alqarqaz M, Koneru J, Mahan M and Ananthasubramaniam K. Applicability, limitations and downstream impact of echocardiography utilization based on the Appropriateness Use Criteria for transthoracic and transesophageal echocardiography. *Int J Cardiovasc Imaging*. 2012;28:1951-8.

40. Bailey SA, Mosteanu I, Tietjen PA, Petrini JR, Alexander J and Keller AM. The use of transthoracic echocardiography and adherence to appropriate use criteria at a regional hospital. *J Am Soc Echocardiogr*. 2012;25:1015-22.

41. Ballo P, Bandini F, Capecchi I, Chiodi L, Ferro G, Fortini A, Giuliani G, Landini G, Laureano R, Milli M, Nenci G, Pizzarelli F, Santoro GM, Vannelli P, Cappelletti C, Zuppiroli A, American College of Cardiology F and American Society of E. Application of 2011 American College of Cardiology Foundation/American Society of Echocardiography appropriateness use criteria in hospitalized patients referred for transthoracic echocardiography in a community setting. *J Am Soc Echocardiogr*. 2012;25:589-98.

42. Bhatia RS, Carne DM, Picard MH and Weiner RB. Comparison of the 2007 and 2011 Appropriate Use Criteria for Transthoracic Echocardiography in Various Clinical Settings. *J Am Soc Echocardiogr*. 2012;25:1162-9.

43. Bhatia RS, Carne DM, Picard MH and Weiner RB. Comparison of the 2007 and 2011 appropriate use criteria for transesophageal echocardiography. *J Am Soc Echocardiogr*. 2012;25:1170-5.

44. Bhatia RS, Kumar V, Picard MH and Weiner RB. Comparison of the 2008 and 2011 appropriate use criteria for stress echocardiography. *J Am Soc Echocardiogr*. 2013;26:339-43.

45. Bhave NM, Mansour IN, Veronesi F, Razi RR, Lang RM and Ward RP. Use of a web-based application of the American College of Cardiology Foundation/American Society of

Echocardiography Appropriateness Use Criteria for Transthoracic Echocardiography: a pilot study. *J Am Soc Echocardiogr*. 2011;24:271-6.

46. Dharmarajan L, Hale TM, Velastegui Z, Castillo E and Kanna B. Utility of two-dimensional echocardiography in pregnancy and post-partum period and impact on management in an inner city hospital. *J Perinat Med*. 2009;37:663-8.

47. Ghatak A, Pullatt R, Vyse S and Silverman DI. Appropriateness criteria are an imprecise measure for repeat echocardiograms. *Echocardiography*. 2011;28:131-5.

48. Grewal GK, Klosterman TB, Shrestha K, Yarmohammadi H, Zurick AO, Varr BC, Tang WH, Lindsay BD and Klein AL. Indications for TEE before cardioversion for atrial fibrillation: implications for appropriateness criteria. *JACC Cardiovasc Imaging*. 2012;5:641-8.

49. Kirkpatrick JN, Ky B, Rahmouni HW, Chirinos JA, Farmer SA, Fields AV, Ogbara J, Eberman KM, Ferrari VA, Silvestry FE, Keane MG, Opotowsky AR, Sutton MS and Wiegers SE. Application of appropriateness criteria in outpatient transthoracic echocardiography. *J Am Soc Echocardiogr*. 2009;22:53-9.

50. Mansour IN, Razi RR, Bhave NM and Ward RP. Comparison of the updated 2011 appropriate use criteria for echocardiography to the original criteria for transthoracic, transesophageal, and stress echocardiography. *J Am Soc Echocardiogr*. 2012;25:1153-61.

51. Martin NM and Picard MH. Use and appropriateness of transthoracic echocardiography in an academic medical center: a pilot observational study. *J Am Soc Echocardiogr*. 2009;22:48-52.

52. Ogbara J, Logani S, Ky B, Chirinos JA, Silvestry FE, Eberman K, Moss JD, Ferrari VA, Keane MG, St John Sutton M, Wiegers SE and Kirkpatrick JN. The utility of prescreening transesophageal echocardiograms: a prospective study. *Echocardiography*. 2011;28:767-73.

53. Parikh PB, Asheld J and Kort S. Does the revised appropriate use criteria for echocardiography represent an improvement over the initial criteria? A comparison between the 2011 and the 2007 appropriateness use criteria for echocardiography. *J Am Soc Echocardiogr*. 2012;25:228-33.

54. Patil HR, Coggins TR, Kusnetzky LL and Main ML. Evaluation of appropriate use of transthoracic echocardiography in 1,820 consecutive patients using the 2011 revised appropriate use criteria for echocardiography. *Am J Cardiol*. 2012;109:1814-7.

55. Rahimi AR, York M, Gheewala N, Markson L, Hauser TH and Manning WJ. Trends in outpatient transthoracic echocardiography: impact of appropriateness criteria publication. *Am J Med*. 2011;124:740-6.

56. Rao G, Sajnani N, Kusnetzky LL and Main ML. Appropriate use of transthoracic echocardiography. *Am J Cardiol*. 2010;105:1640-2.

57. Rao GA, Sajnani NV, Kusnetzky LL and Main ML. Appropriate utilization of transesophageal echocardiography. *Am J Cardiol*. 2009;103:727-9.

58. Silverman GP, Vyse S and Silverman DI. Inappropriately ordered echocardiograms are related to socioeconomic status. *Am J Med Qual*. 2012;27:487-93.

59. Ward RP, Mansour IN, Lemieux N, Gera N, Mehta R and Lang RM. Prospective evaluation of the clinical application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Criteria for transthoracic echocardiography. *JACC Cardiovasc Imaging*. 2008;1:663-71.

60. Willens HJ, Gomez-Marin O, Heldman A, Chakko S, Postel C and Hasan T. Adherence to appropriateness criteria for transthoracic echocardiography: comparisons between a regional department of Veterans Affairs health care system and academic practice and between physicians and mid-level providers. *J Am Soc Echocardiogr*. 2009;22:793-9.

61. Willens HJ, Hendel RC, Inhaber FR, Chakko SC, Postel C, Hasan T and Mohammed F. Appropriateness use criteria for transthoracic echocardiography: relationship with radiology benefit managers preauthorization determination and comparison of the new (2010) criteria to the original (2007) criteria. *Am Heart J*. 2011;162:772-9.

62. Bhattacharyya S, Kamperidis V, Shah BN, Roussin I, Chahal N, Li W, Khattar R and Senior R. Clinical utility and prognostic value of appropriateness criteria in stress echocardiography for the evaluation of valvular heart disease. *JACC Cardiovasc Imaging*. 2013;6:987-92.

63. Schmitz L, Mori N, Khandheria BK and Gupta A. Appropriateness criteria for stress echocardiography in patients with acute chest pain: Are we choosing wisely? *International journal of cardiology*. 2013;2:387-388.

64. Yousuf MI. Using experts' opinions through Delphi technique. *Practical assessment, research & evaluation*. 2007;12:1-8.

65. Cheitlin MD, Armstrong WF, Aurigemma GP, Beller GA, Bierman FZ, Davis JL, Douglas PS, Faxon DP, Gillam LD, Kimball TR, Kussmaul WG, Pearlman AS, Philbrick JT, Rakowski H, Thys DM, Antman EM, Smith SC, Jr., Alpert JS, Gregoratos G, Anderson JL, Hiratzka LF, Hunt SA, Fuster V, Jacobs AK, Gibbons RJ, Russell RO, American College of C, American Heart A and American Society of E. ACC/AHA/ASE 2003 guideline update for the clinical application of echocardiography: summary article: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (ACC/AHA/ASE Committee to Update the 1997 Guidelines for the Clinical Application of Echocardiography). *Circulation*. 2003;108:1146-62.

66. Intersocietal Accreditation Commission (IAC) Echocardiography (ICAEL). IAC Standards and Guidelines for Adult Echocardiography Accreditation. . 2013.

67. Commission MPA and Book AD. Health Care Spending and the Medicare Program. Washington, DC: Medicare Payment Advisory Commission; 2013. 2014.

68. Cortigiani L, Bigi R, Bovenzi F, Molinaro S, Picano E and Sicari R. Prognostic implication of appropriateness criteria for pharmacological pharmacologic stress

echocardiography performed in an outpatient clinic. *Circulation: Cardiovascular Imaging*. 2012;5:298-305.

69. Bhattacharyya S, Kamperidis V, Nalin Shah B, Roussin I, Chahal N, Li W, Khattar R and Senior R. Clinical utility and prognostic value of appropriateness criteria in stress echocardiography for the evaluation of valvular heart disease. *JACC: Cardiovascular Imaging*. 2013;6:987-992.

70. Willens HJ, Nelson K and Hendel RC. Appropriate use criteria for stress echocardiography: Impact of updated criteria on appropriateness ratings, correlation with pre-authorization guidelines, and effect of temporal trends and an educational initiative on utilization. *JACC: Cardiovascular Imaging*. 2013;6:297-309.

71. Schmitz L, Mori N, Khandheria BK and Gupta A. Appropriateness criteria for stress echocardiography in patients with acute chest pain: Are we choosing wisely? *International Journal of Cardiology*. 2013;165:387-388.

72. Bhattacharyya S, Kamperidis V, Chahal N, Shah BN, Roussin I, Li W, Khattar R and Senior R. Clinical and prognostic value of stress echocardiography appropriateness criteria for evaluation of coronary artery disease in a tertiary referral centre. *Heart*. 2014;100:370-4.

73. Bhatia RS, Milford CE, Picard MH and Weiner RB. An educational intervention reduces the rate of inappropriate echocardiograms on an inpatient medical service. *JACC: Cardiovascular Imaging*. 2013;6:545-555.

74. Bhave NM, Mansour IN, Veronesi F, Razi RR, Lang RM and Ward RP. Use of a web-based application of the American College of Cardiology Foundation/American Society of Echocardiography appropriateness use criteria for transthoracic echocardiography: A pilot study. *Journal of the American Society of Echocardiography*. 2011;24:271-276.

75. Armstrong W and Eagle KA. Appropriate use criteria in echocardiography: is no change the same as no benefit? *JAMA Intern Med*. 2013;173:1609-10.

76. Australian Government Department of Human Services. Medicare Australia Statistics, Medicare Item Reports https://www.medicareaustralia.gov.au/statistics/mbs_item.shtml . (accessed July 2016).

77. Gibbons RJ, Smith S and Antman E. American college of cardiology/american heart association clinical practice guidelines: Part I - Where do they come from? *Circulation*. 2003;107:2979-2986.

78. Patel MR, White RD, Abbara S, Bluemke DA, Herfkens RJ, Picard M, Shaw LJ, Silver M, Stillman AE, Udelson J, American College of Radiology Appropriateness Criteria C and American College of Cardiology Foundation Appropriate Use Criteria Task F. 2013 ACCF/ACR/ASE/ASNC/SCCT/SCMR appropriate utilization of cardiovascular imaging in heart failure: a joint report of the American College of Radiology Appropriateness Criteria Committee and the American College of Cardiology Foundation Appropriate Use Criteria Task Force. *J Am Coll Cardiol*. 2013;61:2207-31.

79. R Core Team. R: A Language and Environment for Statistical Computing 2013.

80. Australian Institute of Health and Welfare. *Health expenditure Australia 2011–12. Health and welfare expenditure series no. 50*. Cat. no. HWE 59. Canberra: AIHW; 2013.http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129544656 (accessed August 2014).

81. Australian Institute of Health and Welfare. *Health care expenditure on cardiovascular diseases 2008-09.* Cat. no. CVD 65. Canberra: AIHW; 2014 http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129546379 (accessed September 2014).

82. National Health Performance Authority. *Healthy Communities: Australians' experiences with primary health care in 2010–11*. NHPA; 2013 http://www.nhpa.gov.au/internet/nhpa/publishing.nsf/Content/Healthy-Communities-Australians-experiences-with-primary-health-care-2010-11/$file/HC-PatientExperience-2010-2011-Report.pdf (accessed September 2014).

83. Productivity Commission. *Impacts of Advances in Medical Technology in Australia*, Research Report. Melbourne; 2005. http://www.pc.gov.au/__data/assets/pdf_file/0003/17193/medicaltechnology.pdf (accessed September 2014).

84. Medicare Payment Advisory Commission (MedPAC). *Report to the Congress: Medicare Payment Policy.* 2014. http://www.medpac.gov/documents/reports/mar14_entirereport.pdf?sfvrsn=0 (accessed September 2014). .

85. Medicare Payment Advisory Commission. Report to the Congress: Medicare Payment Policy 2013.

86. OECD. *Geographic Variations in Health Care: What Do We Know and What Can Be Done to Improve Health System Performance?* Paris: OECD Publishing; 2014.

87. Australian Commission on Safety and Quality in Health Care and Australian Institute of Health and Welfare. *Exploring Healthcare Variation in Australia: Analyses Resulting from an OECD Study*. Sydney: ACSQHC; 2014. http://www.safetyandquality.gov.au/wp-content/uploads/2014/05/Exploring-Healthcare-Variation-in-Australia-Analyses-Resulting-from-an-OECD-Study.pdf (accessed September 2014).

88. Joseph P. Newhouse, Alan M. Garber, Robin P. Graham, Margaret A. McCoy, Michelle Mancher and Kibria A. *Variation in Health Care Spending: Target Decision Making, Not Geography*: The National Academies Press; 2013.

89. Public Health England. *The NHS Atlas of Variation in Diagnostic Services. Reducing unwarranted variation to increase value and improve quality.* November 2013. http://www.rightcare.nhs.uk/downloads/Right_Care_Diagnostics_Atlas_hi-res.pdf (accessed August 2014).

90. Horvath J. *Review of Medicare Locals. Report to the Minister for Health and Minister for Sport*. 2014. http://www.health.gov.au/internet/main/publishing.nsf/Content/A69978FAABB1225ECA257CD3001810B7/$File/Review-of-Medicare-Locals-may-2014.pdf (accessed September 2014).

Bibliography

91. Australian Government Department of Health. *Medicare Benefits Schedule Book Category 5. Operating from 01 November 2013* http://www.health.gov.au/internet/mbsonline/publishing.nsf/Content/384D46598C2B03FBCA257BFA002418A2/$File/201311-Cat5.pdf . (accessed September 2014).

92. Australian Bureau of Statistics. *Australian Social Trends: The average Australian*. http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features30April+2013 (accessed August 2014).

93. Medicare locals. Statistics, Boundary and Concordance files: ML Population (ERP 30 June 2012) and Census Statistics-October 2013 http://www.medicarelocals.gov.au/internet/medicarelocals/publishing.nsf/Content/digital-boundaries#.VCEB7_mSx8E . (accessed September 2014).

94. Health Workforce Australia. *National Health Workforce Dataset* http://data.hwa.gov.au/webapi/jsf/tableView/crosstabTableView.xhtml . (accessed September 2014).

95. Australian Bureau of Statistics 2011-2012, table generated 19 August 2014 using: Microdata: Australian Health Survey, National Health Survey, 2011-12 cat. no. 4324.0.55.001, datacube: TableBuilder http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4324.0.55.001main+features12011-12.

96. Bonow RO. 2009 ASNC keynote lecture: measuring cost, cost-effectiveness, and quality in cardiovascular imaging. *J Nucl Cardiol*. 2010;17:362-9.

97. Alter DA, Stukel TA and Newman A. Proliferation of cardiac technology in Canada: a challenge to the sustainability of Medicare. *Circulation*. 2006;113:380-7.

98. GAO. MEDICARE PART B IMAGING SERVICES Rapid Spending Growth and Shift to Physician Offices Indicate Need for CMS to Consider Additional Management Practices. Accessed 1 March 2015: http://www.gao.gov/new.items/d08452.pdf 2008.

99. Hendel RC. Utilization management of cardiovascular imaging pre-certification and appropriateness. *JACC Cardiovasc Imaging*. 2008;1:241-8.

100. Whyte S, Dixon S, Faria R, Walker S, Palmer S and Sculpher M. Getting cost-effective technologies into practice: the value of implementation. An application to B-type natriuretic peptide (BNP) testing in diagnosing chronic heart failure. http://www.eepru.org.uk/Value%20of%20Implementation%20BNP%20testing%20-%20025.pdf Accessed March 2015. 2014.

101. (AIHW) AIoHaW. Cardiovascular disease, diabetes and chronic kidney disease: Australian facts: prevalence and incidence. Cardiovascular, diabetes and chronic kidney disease series no. 2. Cat. no. CDK 2. Canberra: AIHW. 2014.

102. Page K, Ford G, Mangan G, Potter J, Wade V and Grenfell R. PM352 Responding to the unequal burden of cardiovascular disease in Australia: The role of a leading Non Government Organisation (NGO) in tackling health inequities. *Global Heart*. 2014;9:e133.

103. Blackmore CC and Medina LS. Evidence-based radiology and the ACR Appropriateness Criteria. *J Am Coll Radiol*. 2006;3:505-9.

104. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Med Care*. 2001;39:II46-54.

105. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL, American College of Cardiology F and American Heart Association Task Force on Practice G. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2013;62:e147-239.

106. Hiratzka LF, Bakris GL, Beckman JA, Bersin RM, Carr VF, Casey DE, Eagle KA, Hermann LK, Isselbacher EM, Kazerooni EA, Kouchoukos NT, Lytle BW, Milewicz DM, Reich DL, Sen S, Shinn JA, Svensson LG and Williams DM. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM Guidelines for the Diagnosis and Management of Patients With Thoracic Aortic Disease: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine. *Circulation*. 2010;121:e266-e369.

107. Galiè N, Hoeper MM, Humbert M, Torbicki A, Vachiery J-L, Barbera JA, Beghetti M, Corris P, Gaine S, Gibbs JS, Gomez-Sanchez MA, Jondeau G, Klepetko W, Opitz C, Peacock A, Rubin L, Zellweger M, Simonneau G, Vahanian A, Auricchio A, Bax J, Ceconi C, Dean V,

Filippatos G, Funck-Brentano C, Hobbs R, Kearney P, McDonagh T, McGregor K, Popescu BA, Reiner Z, Sechtem U, Sirnes PA, Tendera M, Vardas P, Widimsky P, Sechtem U, Al Attar N, Andreotti F, Aschermann M, Asteggiano R, Benza R, Berger R, Bonnet D, Delcroix M, Howard L, Kitsiou AN, Lang I, Maggioni A, Nielsen-Kudsk JE, Park M, Perrone-Filardi P, Price S, Domenech MTS, Vonk-Noordegraaf A and Zamorano JL. Guidelines for the diagnosis and treatment of pulmonary hypertension. *European Heart Journal*. 2009;30:2493-2537.

108. Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galie N, Pruszczyk P, Bengel F, Brady AJ, Ferreira D, Janssens U, Klepetko W, Mayer E, Remy-Jardin M, Bassand JP and Guidelines ESCCfP. Guidelines on the diagnosis and management of acute pulmonary embolism: the Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). *Eur Heart J*. 2008;29:2276-315.

109. Epstein AE, DiMarco JP, Ellenbogen KA, Estes NA, 3rd, Freedman RA, Gettes LS, Gillinov AM, Gregoratos G, Hammill SC, Hayes DL, Hlatky MA, Newby LK, Page RL, Schoenfeld MH, Silka MJ, Stevenson LW, Sweeney MO, Smith SC, Jr., Jacobs AK, Adams CD, Anderson JL, Buller CE, Creager MA, Ettinger SM, Faxon DP, Halperin JL, Hiratzka LF, Hunt SA, Krumholz HM, Kushner FG, Lytle BW, Nishimura RA, Ornato JP, Page RL, Riegel B, Tarkington LG, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice G, American Association for Thoracic S and Society of Thoracic S. ACC/AHA/HRS 2008 Guidelines for Device-Based Therapy of Cardiac Rhythm Abnormalities: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the ACC/AHA/NASPE 2002 Guideline Update for Implantation of Cardiac Pacemakers and Antiarrhythmia Devices)

developed in collaboration with the American Association for Thoracic Surgery and Society of Thoracic Surgeons. *J Am Coll Cardiol*. 2008;51:e1-62.

110. Task Force for the D, Management of S, European Society of C, European Heart Rhythm A, Heart Failure A, Heart Rhythm S, Moya A, Sutton R, Ammirati F, Blanc JJ, Brignole M, Dahm JB, Deharo JC, Gajek J, Gjesdal K, Krahn A, Massin M, Pepi M, Pezawas T, Ruiz Granell R, Sarasin F, Ungar A, van Dijk JG, Walma EP and Wieling W. Guidelines for the diagnosis and management of syncope (version 2009). *Eur Heart J*. 2009;30:2631-71.

111. Fleisher LA, Beckman JA, Brown KA, Calkins H, Chaikof EL, Fleischmann KE, Freeman WK, Froehlich JB, Kasper EK, Kersten JR, Riegel B, Robb JF, Smith SC, Jr., Jacobs AK, Adams CD, Anderson JL, Antman EM, Buller CE, Creager MA, Ettinger SM, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Hunt SA, Lytle BW, Nishimura R, Ornato JP, Page RL, Riegel B, Tarkington LG, Yancy CW, American College of C, American Heart Association Task Force on Practice G, American Society of E, American Society of Nuclear C, Heart Rhythm S, Society of Cardiovascular A, Society for Cardiovascular A, Interventions, Society for Vascular M, Biology and Society for Vascular S. ACC/AHA 2007 guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 2002 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery) developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, and Society for Vascular Surgery. *J Am Coll Cardiol*. 2007;50:e159-241.

112. Costanzo MR, Dipchand A, Starling R, Anderson A, Chan M, Desai S, Fedson S, Fisher P, Gonzales-Stawinski G, Martinelli L, McGiffin D, Smith J, Taylor D, Meiser B, Webber S, Baran D, Carboni M, Dengler T, Feldman D, Frigerio M, Kfoury A, Kim D, Kobashigawa J, Shullo M, Stehlik J, Teuteberg J, Uber P, Zuckermann A, Hunt S, Burch M, Bhat G, Canter C, Chinnock R, Crespo-Leiro M, Delgado R, Dobbels F, Grady K, Kao W, Lamour J, Parry G, Patel J, Pini D, Towbin J, Wolfel G, Delgado D, Eisen H, Goldberg L, Hosenpud J, Johnson M, Keogh A, Lewis C, O'Connell J, Rogers J, Ross H, Russell S, Vanhaecke J, International Society of H and Lung Transplantation G. The International Society of Heart and Lung Transplantation Guidelines for the care of heart transplant recipients. *J Heart Lung Transplant*. 2010;29:914-56.

113. BlomstromLundqvist C. ACC/AHA/ESC guidelines for the management of patients with supraventricular arrhythmias—executive summary A Report of the American College of Cardiology/American HeartAssociation Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines(Writing Committee to Develop Guidelines for the Management of Patients With Supraventricular Arrhythmias)Developed in collaboration with NASPE–Heart Rhythm Society. *European Heart Journal*. 2003;24:1857-1897.

114. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP, 3rd, Guyton RA, O'Gara PT, Ruiz CE, Skubas NJ, Sorajja P, Sundt TM, 3rd, Thomas JD and American College of Cardiology/American Heart Association Task Force on Practice G. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2014;63:e57-185.

115. Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P and Board of Trustees of the Guidelines International N. Guidelines International Network: toward international standards for clinical practice guidelines. *Annals of internal medicine*. 2012;156:525-31.

116. McNulty EJ, Hung YY, Almers LM, Go AS and Yeh RW. Population trends from 2000-2011 in nuclear myocardial perfusion imaging use. *JAMA*. 2014;311:1248-9.

117. Hendel RC, Patel MR, Kramer CM, Poon M, Hendel RC, Carr JC, Gerstad NA, Gillam LD, Hodgson JM, Kim RJ, Kramer CM, Lesser JR, Martin ET, Messer JV, Redberg RF, Rubin GD, Rumsfeld JS, Taylor AJ, Weigold WG, Woodard PK, Brindis RG, Hendel RC, Douglas PS, Peterson ED, Wolk MJ, Allen JM, Patel MR, American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working G, American College of R, Society of Cardiovascular Computed T, Society for Cardiovascular Magnetic R, American Society of Nuclear C, North American Society for Cardiac I, Society for Cardiovascular A, Interventions and Society of Interventional R. ACCF/ACR/SCCT/SCMR/ASNC/NASCI/SCAI/SIR 2006 appropriateness criteria for cardiac computed tomography and cardiac magnetic resonance imaging: a report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group, American College of Radiology, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, American Society of Nuclear Cardiology, North American Society for Cardiac Imaging, Society for Cardiovascular Angiography and Interventions, and Society of Interventional Radiology. *J Am Coll Cardiol*. 2006;48:1475-97.

118. Hendel RC, Berman DS, Di Carli MF, Heidenreich PA, Henkin RE, Pellikka PA, Pohost GM, Williams KA, American College of Cardiology Foundation Appropriate Use Criteria Task F, American Society of Nuclear C, American College of R, American Heart A, American Society of E, Society of Cardiovascular Computed T, Society for Cardiovascular Magnetic R and Society of Nuclear M. ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM 2009 Appropriate Use Criteria for Cardiac Radionuclide Imaging: A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the American Society of Nuclear Cardiology, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the Society of Cardiovascular Computed Tomography, the Society for Cardiovascular Magnetic Resonance, and the Society of Nuclear Medicine. *J Am Coll Cardiol*. 2009;53:2201-29.

119. Taylor AJ, Cerqueira M, Hodgson JM, Mark D, Min J and O'Gara P. ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 Appropriate use criteria for cardiac computed tomography: a report of the American College of Cardiology Foundation appropriate use criteria task force, the society of cardiovascular computed tomography, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the American Society of Nuclear Cardiology, the North American society for cardiovascular imaging, the society for cardiovascular angiography and interventions, and the society for cardiovascular magnetic resonance. *J Am Coll Cardiol*. 2010.

120. Bhatia RS, Milford CE, Picard MH and Weiner RB. An educational intervention reduces the rate of inappropriate echocardiograms on an inpatient medical service. *JACC Cardiovasc Imaging*. 2013;6:545-55.

121. Willens HJ, Nelson K and Hendel RC. Appropriate use criteria for stress echocardiography: impact of updated criteria on appropriateness ratings, correlation with pre-authorization guidelines, and effect of temporal trends and an educational initiative on utilization. *JACC Cardiovasc Imaging*. 2013;6:297-309.

122. Gibbons RJ, Askew JW, Hodge D, Kaping B, Carryer DJ and Miller T. Appropriate use criteria for stress single-photon emission computed tomography sestamibi studies: a quality improvement project. *Circulation*. 2011;123:499-503.

123. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J and Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of internal medicine*. 2009;151:W-65-W-94.

124. McCully RB, Pellikka PA, Hodge DO, Araoz PA, Miller TD and Gibbons RJ. Applicability of appropriateness criteria for stress imaging: similarities and differences between stress echocardiography and single-photon emission computed tomography myocardial perfusion imaging criteria. *Circulation Cardiovascular imaging*. 2009;2:213-8.

125. Mansour IN, Lang RM, Aburuwaida WM, Bhave NM and Ward RP. Evaluation of the clinical application of the ACCF/ASE appropriateness criteria for stress echocardiography. *J Am Soc Echocardiogr*. 2010;23:1199-204.

126. Lin FY, Dunning AM, Narula J, Shaw LJ, Gransar H, Berman DS and Min JK. Impact of an automated multimodality point-of-order decision support tool on rates of appropriate

testing and clinical decision making for individuals with suspected coronary artery disease: a prospective multicenter study. *J Am Coll Cardiol*. 2013;62:308-16.

127. Cortigiani L, Bigi R, Bovenzi F, Molinaro S, Picano E and Sicari R. Prognostic implication of appropriateness criteria for pharmacologic stress echocardiography performed in an outpatient clinic [corrected]. *Circulation Cardiovascular imaging*. 2012;5:298-305.

128. Gibbons RJ, Miller TD, Hodge D, Urban L, Araoz PA, Pellikka P and McCully RB. Application of appropriateness criteria to stress single-photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. *J Am Coll Cardiol*. 2008;51:1283-9.

129. Mehta R, Ward RP, Chandra S, Agarwal R, Williams KA, American College of Cardiology F and American Society of Nuclear C. Evaluation of the American College of Cardiology Foundation/American Society of Nuclear Cardiology appropriateness criteria for SPECT myocardial perfusion imaging. *J Nucl Cardiol*. 2008;15:337-44.

130. Hendel RC, Cerqueira M, Douglas PS, Caruth KC, Allen JM, Jensen NC, Pan W, Brindis R and Wolk M. A multicenter assessment of the use of single-photon emission computed tomography myocardial perfusion imaging with appropriateness criteria. *J Am Coll Cardiol*. 2010;55:156-62.

131. Gibbons RJ, Askew JW, Hodge D and Miller TD. Temporal trends in compliance with appropriateness criteria for stress single-photon emission computed tomography sestamibi studies in an academic medical center. *Am Heart J*. 2010;159:484-9.

132. Carryer DJ, Hodge DO, Miller TD, Askew JW and Gibbons RJ. Application of appropriateness criteria to stress single photon emission computed tomography sestamibi studies: a comparison of the 2009 revised appropriateness criteria to the 2005 original criteria. *Am Heart J*. 2010;160:244-9.

133. Gupta A, Tsiaras SV, Dunsiger SI and Tilkemeier PL. Gender disparity and the appropriateness of myocardial perfusion imaging. *J Nucl Cardiol*. 2011;18:588-94.

134. Gholamrezanezhad A, Shirafkan A, Mirpour S, Rayatnavaz M, Alborzi A, Mogharrabi M, Hassanpour S and Ramezani M. Appropriateness of referrals for single-photon emission computed tomography myocardial perfusion imaging (SPECT-MPI) in a developing community: a comparison between 2005 and 2009 versions of ACCF/ASNC appropriateness criteria. *J Nucl Cardiol*. 2011;18:1044-52.

135. Druz RS, Phillips LM and Sharifova G. Clinical evaluation of the appropriateness use criteria for single-photon emission-computed tomography: differences by patient population, physician specialty, and patient outcomes. *ISRN cardiology*. 2011;2011.

136. Soine LA, Cunningham SL, Motzer SA, Inoue LY and Caldwell JH. Application of appropriate use criteria for stress myocardial perfusion imaging at two academic medical centers: compliance and association with image findings. *J Am Acad Nurse Prac*. 2012;24:200-8.

137. Koh AS, Flores JL, Keng FY, Tan RS and Chua TS. Evaluation of the American College of Cardiology Foundation/American Society of Nuclear Cardiology appropriateness

criteria for SPECT myocardial perfusion imaging in an Asian tertiary cardiac center. *J Nucl Cardiol*. 2011;18:324-30.

138. Nelson KH, Willens HJ and Hendel RC. Utilization of radionuclide myocardial perfusion imaging in two health care systems: assessment with the 2009 ACCF/ASNC/AHA appropriateness use criteria. *J Nucl Cardiol*. 2012;19:37-42.

139. Koh AS, Flores JL, Keng FY, Tan RS and Chua TS. Correlation between clinical outcomes and appropriateness grading for referral to myocardial perfusion imaging for preoperative evaluation prior to non-cardiac surgery. *J Nucl Cardiol*. 2012;19:277-84.

140. Winchester DE, Meral R, Ryals S, Beyth RJ and Shaw LJ. Appropriate use of myocardial perfusion imaging in a veteran population: profit motives and professional liability concerns. *JAMA Intern Med*. 2013;173:1381-3.

141. Doukky R, Hayes K, Frogge N, Balakrishnan G, Dontaraju VS, Rangel MO, Golzar Y, Garcia-Sayan E and Hendel RC. The impact of appropriate use on the prognostic value of SPECT myocardial perfusion imaging. *Circulation*. 2013:CIRCULATIONAHA. 113.002744.

142. Moralidis E, Papadimitriou N, Stathaki M, Xourgia X, Spyridonidis T, Fotopoulos A, Apostolopoulos D, Karkavitsas N and Gotzamani-Psarrakou A. A multicenter evaluation of the appropriate use of single-photon emission tomography myocardial perfusion imaging in Greece. *J Nucl Cardiol*. 2013;20:275-83.

143. Aldweib N, Negishi K, Seicean S, Jaber WA, Hachamovitch R, Cerqueira M and Marwick TH. Appropriate test selection for single-photon emission computed tomography

imaging: association with clinical risk, posttest management, and outcomes. *Am Heart J*. 2013;166:581-8.

144. Ayyad AE, Cole J, Syed A, Desai MY, Halliburton S, Schoenhagen P, Flamm SD and Sola S. Temporal trends in utilization of cardiac computed tomography. *J Cardiovasc Comput Tomogr*. 2009;3:16-21.

145. Miller JA, Raichlin E, Williamson EE, McCully RB, Pellikka PA, Hodge DO, Miller TD, Gibbons RJ and Araoz PA. Evaluation of coronary CTA Appropriateness Criteria in an academic medical center. *J Am Coll Radiol*. 2010;7:125-31.

146. Murphy MK, Brady TJ, Nasir K, Gazelle GS, Bamberg F, Truong QA, Mamuya WS, Abbara S, Lee TH and Blankstein R. Appropriateness and utilization of cardiac CT: implications for development of future criteria. *J Nucl Cardiol*. 2010;17:881-9.

147. El Sibai K, Itani S, Rabah A, Hourani M and Dakik HA. Evaluation of the appropriateness criteria for coronary computed tomography angiography in an academic medical center in a developing country: comparison of the 2006 and 2010 criteria. *J Nucl Cardiol*. 2011;18:1053-8.

148. Chinnaiyan KM, Peyser P, Goraya T, Ananthasubramaniam K, Gallagher M, Depetris A, Boura JA, Kazerooni E, Poopat C, Al-Mallah M, Saba S, Patel S, Girard S, Song T, Share D and Raff G. Impact of a continuous quality improvement initiative on appropriate use of coronary computed tomography angiography. Results from a multicenter, statewide registry, the Advanced Cardiovascular Imaging Consortium. *J Am Coll Cardiol*. 2012;60:1185-91.

149. Rich ME, Utsunomiya D, Simprini LA, Weigold WG, Weissman G and Taylor AJ. Prospective evaluation of the updated 2010 ACCF cardiac CT appropriate use criteria. *Journal of cardiovascular computed tomography*. 2012;6:108-112.

150. Mazimba S, Grant N, Parikh A, Patel T, Dahale B, Franco Z, Dittoe N, Shah T and Hahn HS. Comparison of the 2006 and 2010 cardiac CT appropriateness criteria in a real-world setting. *J Am Coll Radiol*. 2012;9:630-4.

151. Wasfy MM, Brady TJ, Abbara S, Nasir K, Hoffmann U, Cury RC, Di Carli MF and Blankstein R. Comparison of cardiac computed tomography examination appropriateness under the 2010 revised versus the 2006 original Appropriate Use Criteria. *J Cardiovasc Comput Tomogr*. 2012;6:99-107.

152. Cullen EL, Aggarwal SR, Goss BC, Hodge DO, Gibbons RJ and Araoz PA. Comparison of the applicability of the 2006 and the 2010 cardiac CT angiography appropriate use criteria. *J Am Coll Radiol*. 2013;10:258-67.

153. Davis DA and TaylorVaisey A. Translating guidelines into practice - A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *Can Med Assoc J*. 1997;157:408-416.

154. Cook C, Cole G, Asaria P, Jabbour R and Francis DP. The annual global economic burden of heart failure. *Int J Cardiol*. 2014;171:368-76.

155. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA, Maddox TM, Nichol G, Pham M, Pina IL, Trogdon

JG, American Heart Association Advocacy Coordinating C, Council on Arteriosclerosis T, Vascular B, Council on Cardiovascular R, Intervention, Council on Clinical C, Council on E, Prevention and Stroke C. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail*. 2013;6:606-19.

156. Krum H and Abraham WT. Heart failure. *Lancet*. 2009;373:941-55.

157. Page K, Marwick TH, Lee R, Grenfell R, Abhayaratna WP, Aggarwal A, Briffa TG, Cameron J, Davidson PM, Driscoll A, Garton-Smith J, Gascard DJ, Hickey A, Korczyk D, Mitchell JA, Sanders R, Spicer D, Stewart S, Wade V and National Heart Foundation of A. A systematic approach to chronic heart failure care: a consensus statement. *Med J Aust*. 2014;201:146-50.

158. Chiriac A, Kadkhodayan A, Pislaru SV, Dailey EJ, Scott CG, Buechler TE, Newman JS and Pellikka PA. Clinical Importance of Transthoracic Echocardiography with Direct Input from Treating Physicians. *Journal of the American Society of Echocardiography*.

159. Pocock SJ, Ariti CA, McMurray JJ, Maggioni A, Kober L, Squire IB, Swedberg K, Dobson J, Poppe KK, Whalley GA, Doughty RN and Meta-Analysis Global Group in Chronic Heart F. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J*. 2013;34:1404-13.

160. Hosmer DW, Lemeshow S and May S. Model Development *Applied Survival Analysis*: John Wiley & Sons, Inc.; 2008: 132-168.

161. Institute UCaTS. Sample size – Survival analysis. 2016.

162. Di Lenarda A, Sabbadini G, Pinamonti B and Sinagra G. [Echocardiographic follow-up of patients with heart failure. Which parameters should be measured? How often?]. *Italian heart journal Supplement : official journal of the Italian Federation of Cardiology*. 2000;1:1404-10.

163. Medicare Payment Advisory Commission (MedPAC). Report to the Congress: Medicare Payment Policy 2013. URL http://www.medpac.gov/documents/reports/mar13_entirereport.pdf?sfvrsn=0 Accessed 29 June 2015.

164. Medicare Payment Advisory Commission (MedPAC). Report to the Congress: Medicare Payment Policy 2005. URL http://www.medpac.gov/documents/Mar05_EntireReport.pdf Accessed 28 May 2014.

165. Gurzun M-M and Ionescu A. Appropriateness of use criteria for transthoracic echocardiography: are they relevant outside the USA? *European Heart Journal - Cardiovascular Imaging*. 2013.

166. Sachdeva R and Douglas PS. Quality Improvement Interventions to Improve Appropriateness of Imaging Studies: Necessary, But Are They Sufficient? *Circulation: Cardiovascular Quality and Outcomes*. 2016.

167. Chaudhuri D, Montgomery A, Gulenchyn K, Mitchell M and Joseph P. Effectiveness of Quality Improvement Interventions at Reducing Inappropriate Cardiac Imaging: A

Systematic Review and Meta-Analysis. *Circulation: Cardiovascular Quality and Outcomes*. 2016.

168. Massari F, Iacoviello M, Scicchitano P, Mastropasqua F, Guida P, Riccioni G, Speziale G, Caldarola P, Ciccone MM and Di Somma S. Accuracy of bioimpedance vector analysis and brain natriuretic peptide in detection of peripheral edema in acute and chronic heart failure. *Heart & lung : the journal of critical care*. 2016;45:319-26.

169. Ciccone MM, Aquilino A, Cortese F, Scicchitano P, Sassara M, Mola E, Rollo R, Caldarola P, Giorgino F, Pomo V and Bux F. Feasibility and effectiveness of a disease and care management model in the primary health care system for patients with heart failure and diabetes (Project Leonardo). *Vascular health and risk management*. 2010;6:297-305.

170. Chaudhuri D, Montgomery A, Gulenchyn K, Mitchell M and Joseph P. Effectiveness of Quality Improvement Interventions at Reducing Inappropriate Cardiac Imaging: A Systematic Review and Meta-Analysis. *Circulation Cardiovascular quality and outcomes*. 2016;9:7-13.

171. Hussey PS, Timbie JW, Burgette LF, Wenger NS, Nyweide DJ and Kahn KL. Appropriateness of advanced diagnostic imaging ordering before and after implementation of clinical decision support systems. *JAMA*. 2015;313:2181-2.

172. Dudzinski DM, Bhatia RS, Mi MY, Isselbacher EM, Picard MH and Weiner RB. Effect of Educational Intervention on the Rate of Rarely Appropriate Outpatient Echocardiograms

Ordered by Attending Academic Cardiologists: A Randomized Clinical Trial. *JAMA Cardiol*. 2016;1:805-812.

173. Bhatia RS, Alabousi M, Dudzinski DM and Weiner RB. Appropriate use criteria: a review of need, development and applications. *Expert Rev Cardiovasc Ther*. 2016;14:281-90.

174. De Nardo A and Niranjan S. Utilisation of echocardiography and application of the appropriate use criteria at a large tertiary hospital in Queensland. *Australasian Journal of Ultrasound in Medicine*. 2016;19:64-70.

175. Patton MQ. Qualitative Research *Encyclopedia of Statistics in Behavioral Science*: John Wiley & Sons, Ltd; 2005.

176. Corbin J and Strauss A. Grounded Theory Research - Procedures, Canons and Evaluative Criteria. *Z Soziol*. 1990;19:418-427.

177. Ezzy D. *Qualitative analysis*: Routledge; 2013.

178. Saini V, Garcia-Armesto S, Klemperer D, Paris V, Elshaug AG, Brownlee S, Ioannidis JP and Fisher ES. Drivers of poor medical care. *Lancet*. 2017.

179. Simpkin AL and Schwartzstein RM. Tolerating Uncertainty - The Next Medical Revolution? *N Engl J Med*. 2016;375:1713-1715.

180. Glasziou P, Straus S, Brownlee S, Trevena L, Dans L, Guyatt G, Elshaug AG, Janett R and Saini V. Evidence for underuse of effective medical services around the world. *Lancet*. 2017.

181. Fonseca R, Pathan F and Marwick T. Identification of Rarely Appropriate Echocardiograms in the Echocardiography Laboratory. *Journal of the American College of Cardiology*. 2016;67:1603.

182. Ward RP, Krauss D, Mansour IN, Lemieux N, Gera N, Lang RM, American College of Cardiology F and American Society of E. Comparison of the clinical application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Criteria for outpatient transthoracic echocardiography in academic and community practice settings. *J Am Soc Echocardiogr*. 2009;22:1375-81.

183. Lin GA and Harris IS. Can appropriate use criteria lead to appropriate care? *Circulation Cardiovascular quality and outcomes*. 2015;8:4-5.

184. Goldzweig CL, Orshansky G, Paige NM, Miake-Lye IM, Beroes JM, Ewing BA and Shekelle PG. Electronic health record-based interventions for improving appropriate diagnostic imaging: a systematic review and meta-analysis. *Annals of internal medicine*. 2015;162:557-65.

185. Levitt K, Shojania KG and Bhatia RS. Point-of-care decision support for reducing inappropriate test use: easier said than done. *BMJ quality & safety*. 2016;25:6-8.

186. Boggan JC, Schulteis RD, Donahue M and Simel DL. Guideline-based decision support has a small, non-sustained effect on transthoracic echocardiography ordering frequency. *BMJ quality & safety*. 2015.

187. Bhattacharyya S and Lloyd G. Improving Appropriateness and Quality in Cardiovascular Imaging: A Review of the Evidence. *Circulation Cardiovascular imaging*. 2015;8.

188. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF and Group S. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.

189. Dehmer GJ and Patel MR. The Use of Appropriate Use Criteria Are Increasing, But What Are Their Effects on Medical Care? *Circulation*. 2015.

190. Shor R. Using Appropriate Use Criteria to Address Pre-Authorization. *J Am Coll Cardiol*. 2015;66:1300-2.

191. Banihashemi B, Maftoon K, Chow BJ, Bernick J, Wells GA and Burwash IG. Limitations of free-form-text diagnostic requisitions as a tool for evaluating adherence to appropriate use criteria for transthoracic echocardiography. *Cardiovasc Ultrasoun*. 2015;13:4.

192. Al-Kaisey A, Jones E, Nadurata V, Farouque O, De Silva D and Ramchand J. Appropriate use of echocardiography in an Australian regional centre. *Intern Med J*. 2015;45:1128-33.

193. Cassel CK and Guest JA. Choosing wisely: helping physicians and patients make smart decisions about their care. *JAMA*. 2012;307:1801-2.

194. Khan HA, Wineinger NE, Uddin PQ, Mehta HS, Rubenson DS and Topol EJ. Can hospital rounds with pocket ultrasound by cardiologists reduce standard echocardiography? *Am J Med*. 2014;127:669 e1-7.

195. Mehta M, Jacobson T, Peters D, Le E, Chadderdon S, Allen AJ, Caughey AB and Kaul S. Handheld ultrasound versus physical examination in patients referred for transthoracic echocardiography for a suspected cardiac condition. *JACC Cardiovasc Imaging*. 2014;7:983-90.

196. Vourvouri EC, Koroleva LY, Ten Cate FJ, Poldermans D, Schinkel AF, van Domburg RT, Vletter WB and Roelandt JR. Clinical utility and cost effectiveness of a personal ultrasound imager for cardiac evaluation during consultation rounds in patients with suspected cardiac disease. *Heart*. 2003;89:727-30.

197. Greaves K, Jeetley P, Hickman M, Dwivedi G, Sabharwal N, Lim T, Janardhanan R and Senior R. The use of hand-carried ultrasound in the hospital setting--a cost-effective analysis. *J Am Soc Echocardiogr*. 2005;18:620-5.

198. Galasko GI, Barnes SC, Collinson P, Lahiri A and Senior R. What is the most cost-effective strategy to screen for left ventricular systolic dysfunction: natriuretic peptides, the electrocardiogram, hand-held echocardiography, traditional echocardiography, or their combination? *Eur Heart J*. 2006;27:193-200.

199. Trambaiolo P, Papetti F, Posteraro A, Amici E, Piccoli M, Cerquetani E, Pastena G, Gambelli G and Salustri A. A hand-carried cardiac ultrasound device in the outpatient cardiology clinic reduces the need for standard echocardiography. *Heart*. 2007;93:470-5.

200. Gianstefani S, Catibog N, Whittaker AR, Ioannidis AG, Vecchio F, Wathen PT, Douiri A, Reiken J and Monaghan MJ. Pocket-size imaging device: effectiveness for ward-based transthoracic studies. *Eur Heart J Cardiovasc Imaging*. 2013;14:1132-9.

201. Kitada R, Fukuda S, Watanabe H, Oe H, Abe Y, Yoshiyama M, Song JM, Sitges M, Shiota T, Ito H and Yoshikawa J. Diagnostic accuracy and cost-effectiveness of a pocket-sized transthoracic echocardiographic imaging device. *Clin Cardiol*. 2013;36:603-10.

202. Kini V, Mehta N, Mazurek JA, Ferrari VA, Epstein AJ, Groeneveld PW and Kirkpatrick JN. Focused Cardiac Ultrasound in Place of Repeat Echocardiography: Reliability and Cost Implications. *J Am Soc Echocardiogr*. 2015;28:1053-9.

203. Gurzun M-M and Ionescu A. Appropriateness of use criteria for transthoracic echocardiography: are they relevant outside the USA? *Eur Heart J-Card Img*. 2013:jet186.

204. Ward RP. The role of appropriateness criteria in the clinical practice of echocardiography. *curr cardiovasc imaging rep*. 2009;2:319-324.

205. Centers for Disease Control and Prevention. National Hospital Discharge Survey: 2010 table, Number and rate of hospital discharges. 2010.

206. Spencer KT, Kimura BJ, Korcarz CE, Pellikka PA, Rahko PS and Siegel RJ. Focused cardiac ultrasound: recommendations from the American Society of Echocardiography. *J Am Soc Echocardiogr*. 2013;26:567-81.

Appendix

Appendix Table 1 . Population, number of doctors and burden of cardiovascular disease in Australian Medicare Locals

Medicare Local	Peer group	State	Region code	Population 2012	Mean age	%Women 2012	% ≥65years old	Number Practitioners	Non-cardiologists per 1,000 pop	Cardiologists per 100,000 pop	CV deaths per 1,000 pop	%population>18 years with Cardiovascular diseases	% Peopl e living near big cities	%Least disadvantage people
Eastern Sydney	Metro 1	NSW	ML101	388696	38.1	49.7	12.5	3779	9.58	13.72	1.2	16.16	99.9	66.2
Inner West Sydney	Metro 1	NSW	ML102	590233	37.6	50.0	11.9	2860	4.75	9.37	1.4	19.1	100.0	44.3
South Eastern Sydney	Metro 2	NSW	ML103	469806	39.1	50.7	15.1	1646	3.45	4.78	2.1	22.6	100.0	51.3
South Western Sydney	Metro 3	NSW	ML104	887950	36.4	50.4	11.9	2499	2.78	3.36	1.7	16.0	100.0	21.3
Western Sydney	Metro 3	NSW	ML105	862669	35.6	49.8	10.7	3299	3.77	5.37	1.2	21.1	100.0	40.0

Nepean - Blue Mountains	Regional 1	NSW	ML106	351262	37.3	50.3	12.3	1085	3.05	3.20	1.6	27.8	99.8	38.8
Northern Sydney	Metro 1	NSW	ML107	404957	39.3	51.2	15.5	1725	4.21	4.81	2.2	12.4	100.0	83.9
Sydney North Shore and Beaches	Metro 1	NSW	ML108	459529	39.1	51.1	14.7	2516	5.4	7.20	2.7	18.5	100.0	88.9
Central Coast NSW	Regional 1	NSW	ML109	325295	40.8	51.4	19.2	1028	3.13	2.83	3.1	27.8	100.0	28.8
Illawarra - Shoalhaven	Regional 1	NSW	ML110	387558	40.4	50.2	18.2	1140	2.91	2.92	2.8	26.1	99.9	28.5
Hunter	Regional 1	NSW	ML111	701695	40.1	50.2	17.6	2526	3.56	3.66	2.5	26.7	98.6	27.2
North Coast NSW	Regional 2	NSW	ML113	497721	42.6	50.9	20.7	1583	3.16	1.86	3.4	25.0	87.5	12.9
New England	Regional 2	NSW	ML114	183341	39.6	50.3	17.4	440	2.40	0.00	2.9	23.2	39.8	20.6
Western NSW	Regional 2	NSW	ML115	256476	39.0	49.9	16.8	685	2.67	2.00	2.5	29.9	61.3	24.3
Murrumbidgee	Regional 2	NSW	ML116	185911	38.9	49.7	16.4	437	2.35	3.28	3.7	22.5	56.4	25.5
Southern NSW	Regional 2	NSW	ML117	198363	41.2	49.7	17.9	377	1.9	0.00	2.9	21.8	70.6	31.7
Far West NSW	Rural 1	NSW	ML118	38530	39.8	48.9	16.9	100	2.6	0.00	2.7	27.0	0.0	5.6

Appendix

Inner North West Melbourne	Metro 1	VIC	ML201	448879	37.2	50.7	12.1	4778	10.49	14.66	1.0	13.3	100.0	50.3
Bayside	Metro 1	VIC	ML202	592709	39.5	51.1	14.9	2849	4.72	7.82	1.8	23.7	100.0	79.8
South Western Melbourne	Metro 2	VIC	ML203	267603	34.6	50.0	9.1	435	1.61	1.12	3.2	21.6	100.0	36.2
Macedon Ranges and North Western Melbourne	Metro 3	VIC	ML204	487218	35.9	49.9	10.5	1202	2.44	2.33	0.9	13.4	100.0	28.9
Northern Melbourne	Metro 3	VIC	ML205	641238	36.8	50.6	12.2	2226	3.42	5.00	1.8	19.9	100.0	40.2
Inner East Melbourne	Metro 1	VIC	ML206	624018	40.1	51.2	17.2	3073	4.84	8.14	2.1	19.7	100.0	72.7
Eastern Melbourne	Metro 2	VIC	ML207	411822	38.5	50.7	13.6	915	2.2	1.47	1.9	28.2	100.0	57.1
South Eastern Melbourne	3. Metro 3	VIC	ML208	492648	35.6	49.9	10.9	1019	2.05	1.25	1.4	20.6	100.0	26.4

Frankston - Mornington Peninsula	Regional 1	VIC	ML209	282319	40.6	51.2	18.3	879	3.08	3.29	2.6	24.2	100.0	39.7
Barwon	Regional 1	VIC	ML210	280892	40	50.5	16.9	982	3.45	3.70	2.6	26.5	99.7	36.2
Grampians	Regional 2	VIC	ML211	209600	40.6	50.4	17.9	540	2.54	3.46	3.0	34.9	79.0	22.1
Great South Coast	Regional 2	VIC	ML212	101752	40.5	50.1	17.9	246	2.41	0.00	3.4	16.3	65.0	23.5
Lower Murray	Rural 1	VIC	ML213	67141	39.3	49.9	16.5	181	2.65	4.47	2.0	28.7	0.0	11.5
Loddon - Mallee - Murray	Regional 2	VIC	ML214	218299	40.9	50.2	18.6	569	2.57	3.33	2.9	23.9	79.7	20.1
Goulburn Valley	Regional 2	VIC	ML215	150360	39.7	49.6	16.7	366	2.43	0.00	3.5	15.9	98.4	20.1
Hume	Regional 2	VIC	ML216	201483	40.6	50.3	18	547	2.69	2.03	2.4	26.4	86.2	26.1
Gippsland	Regional 2	VIC	ML217	262285	41.4	50.2	19.1	615	2.33	1.14	2.8	24.5	80.6	21.1
Metro North Brisbane	Metro 2	QLD	ML301	907533	37.2	50.4	12.7	4962	5.37	9.52	1.9	17.5	99.9	56.4
Greater Metro South Brisbane	Metro 2	QLD	ML302	932531	36.4	50.2	11.5	3625	3.84	4.41	1.8	22.8	99.4	47.2
Gold Coast	Metro 2	QLD	ML303	539783	38.7	50.8	14.7	1842	3.36	5.04	1.7	20.4	100.0	39.2

Sunshine Coast	Regional 1	QLD	ML304	371780	41.2	51.1	18.6	1335	3.53	5.21	2.7	21.6	99.4	28.2
West Moreton - Oxley	Metro 3	QLD	ML305	370234	35.6	49.9	11.2	769	2.07	0.83	1.4	20.0	100.0	29.5
Darling Downs - South West Queensland	Regional 2	QLD	ML306	297513	38.5	50.3	15.9	793	2.65	1.24	2.2	20.6	70.3	23.9
Wide Bay	Regional 2	QLD	ML307	208583	42.2	50.5	20.9	673	3.2	1.69	2.4	22.7	90.0	6.1
Central Queensland	Rural 1	QLD	ML308	216470	36.1	48.5	11.4	593	2.72	1.71	1.2	22.0	75.9	34.2
Central and North West Queensland	Rural 2	QLD	ML309	44865	34.2	46.9	8.8	125	2.78	0.00	1.8	_	0.0	21.8
Townsville - Mackay	Rural 1	QLD	ML310	409950	36.1	48.6	11	1414	3.41	3.57	1.7	24.2	19.8	33.5
Far North Queensland	Rural 2	QLD	ML311	266286	36.9	49.5	11.9	903	3.36	2.94	1.8	18.9	0.0	23.9
Northern Adelaide	Metro 3	SA	ML401	405065	37.6	50.2	13.6	976	2.4	1.03	2.0	31.1	100.0	22.6

Central Adelaide and Hills	Metro 2	SA	ML402	517032	40.5	51.0	16.9	3709	7.05	11.57	2.9	22.2	100.0	43.5
Southern Adelaide - Fleurieu - Kangaroo Island	Metro 2	SA	ML403	401464	40.9	51.2	17.7	1501	3.67	6.48	2.6	25.0	98.9	41.4
Country South SA	Regional 2	SA	ML404	133718	40.6	49.2	17.5	255	1.9	0.00	2.6	29.8	21.3	13.1
Country North SA	Rural 1	SA	ML405	199020	40.5	48.9	17.5	312	1.56	0.00	2.9	23.9	23.1	18.5
Perth Central and East Metro	Metro 2	WA	ML501	470311	37.8	49.7	12.7	4235	8.89	10.69	1.6	17.0	100.0	57.5
Perth North Metro	Metro 2	WA	ML502	531958	36.8	50.3	11.8	985	1.84	0.96	1.3	18.4	100.0	60.9
Fremantle	Metro 2	WA	ML503	240394	38.6	50.7	13.7	1114	4.58	5.19	2.6	10.8	99.9	66.7
Bentley - Armadale	Metro 2	WA	ML504	418802	36.5	49.5	11.4	692	1.65	0.00	1.6	15.6	100.0	44.8

Appendix

Perth South Coastal	Regional 1	WA	ML505	243056	37.6	50.2	14.7	472	1.93	1.23	0.9	23.3	100.0	34.0
South West WA	Regional 2	WA	ML506	299005	39.2	50.0	15.2	639	2.13	0.00	2.0	28.5	59.2	29.8
Goldfields - Midwest	Rural 2	WA	ML507	127272	36.3	47.8	10.5	256	2.01	0.00	1.5	19.7	0.0	28.3
Kimberley - Pilbara	Rural 2	WA	ML508	101907	32.5	41.4	3.2	285	2.79	0.00	0.6	20.7	0.0	37.8
Tasmania	Regional 2	TAS	ML601	512333	40.2	50.1	16.8	1970	3.81	3.36	2.7	29.1	65.6	21.5
Northern Territory	Rural 2	NT	ML701	235182	32.9	47.5	5.9	1068	4.52	1.28	0.9	14.7	0.0	37.1
Australian Capital Territory	Metro 1	ACT	ML801	374909	36.4	50.2	11	1747	4.62	4.00	1.5	23.1	100.0	78.5

Appendix Table 2 . Rates of cardiac imaging among Medicare Locals

Medicare Local	**Peer group**	**State**	**Region**	**TTE per 100,000 pop (age-weighted)**	**TTE rank**	**TOE per 100,000 pop (age-weighted)**	**TOE rank**	**SE per 100,000 pop (age-weighted)**	**SE rank**	**SPECT per 100,000 pop (age-weighted)**	**SPECT rank**
Central Adelaide and Hills	2. Metro 2	SA	ML402	7184.1	1	120.2	7	878.7	18	263.3	27
Inner North West Melbourne	1. Metro 1	VIC	ML201	6894.0	2	197.1	2	2094	5	685.5	7
Eastern Sydney	1. Metro 1	NSW	ML101	6831.4	3	242.1	1	2852	2	316.7	24
Wide Bay	5. Regional 2	QLD	ML307	6072.6	4	0		0		–	
Inner East Melbourne	1. Metro 1	VIC	ML206	5476.3	5	76.4	12	2366	3	262.2	28
Fremantle	2. Metro 2	WA	ML503	5281.0	6	97.1	10	1821	6	131.8	35
Inner West Sydney	1. Metro 1	NSW	ML102	5129.5	7	70.5	15	2105	4	389	19
Perth Central and East Metro	2. Metro 2	WA	ML501	4912.0	8	143.3	4	976.7	16	313.9	25

Sunshine Coast	4. Regional 1	QLD	ML304	4874.8	9	0		420.2	30	805.7	3
Central Coast NSW	4. Regional 1	NSW	ML109	4659.7	10	73.6	14	519	26	796.1	4
Frankston - Mornington Peninsula	4. Regional 1	VIC	ML209	4583.8	11	0		264.9	33	531.7	9
Gold Coast	2. Metro 2	QLD	ML303	4538.4	12	108.5	9	1142	12	239.4	33
Metro North Brisbane	2. Metro 2	QLD	ML301	4528.3	13	134.5	6	1115	13	442.5	15
Loddon - Mallee - Murray	5. Regional 2	VIC	ML214	4461.2	14	55.2	22	0		259.6	29
Grampians	5. Regional 2	VIC	ML211	4352.3	15	44.5	27	538.3	24	789.8	5
Western Sydney	3. Metro 3	NSW	ML105	4277.5	16	56.7	19	1637	10	244.8	31
Sydney North Shore and Beaches	1. Metro 1	NSW	ML108	4202.4	17	116.2	8	1674	9	438.6	16
South Eastern Sydney	2. Metro 2	NSW	ML103	4138.8	18	136.6	5	2992	1	371	20
Bayside	1. Metro 1	VIC	ML202	4024.5	19	43.6	28	827.7	20	453.1	13

Greater Metro South Brisbane	2. Metro 2	QLD	ML302	3908.6	20	80.1	11	1106	15	298.5	26
Western NSW	5. Regional 2	NSW	ML115	3752.8	21	56.4	20	519.6	25	930.2	2
Hume	5. Regional 2	VIC	ML216	3643.7	22	0		886.7	17	641.8	8
North Coast NSW	5. Regional 2	NSW	ML113	3550.4	23	48.1	25	1817	7	240.5	32
Northern Sydney	1. Metro 1	NSW	ML107	3542.3	24	170.6	3	1334	11	453	14
Southern Adelaide - Fleurieu - Kangaroo Island	2. Metro 2	SA	ML403	3453.6	25	46.3	26	830	19	–	
Lower Murray	6. Rural 1	VIC	ML213	3276.1	26	_		0		–	
Illawarra - Shoalhaven	4. Regional 1	NSW	ML110	3210.3	27	67.4	16	762.6	21	955.8	1
South Western Sydney	3. Metro 3	NSW	ML104	3146.6	28	56.4	21	1114	14	461.9	12
Hunter	4. Regional 1	NSW	ML111	3141.4	29	34	30	719.9	23	422.8	17
Townsville - Mackay	6. Rural 1	QLD	ML310	3059.7	30	54.4	23	0		329.1	23

Nepean - Blue Mountains	4. Regional 1	NSW	ML106	2984.0	31	61.6	17	508	27	356.6	21
Barwon	4. Regional 1	VIC	ML210	2729.2	32	0		0		–	
Australian Capital Territory	1. Metro 1	ACT	ML801	2611.2	33	0		253	34	253.8	30
Eastern Melbourne	2. Metro 2	VIC	ML207	2502.3	34	35.2	29	752.2	22	341.2	22
Northern Melbourne	3. Metro 3	VIC	ML205	2454.5	35	60.4	18	503	28	529	10
Darling Downs - South West Queensland	5. Regional 2	QLD	ML306	2328.7	36	_		0	37	–	
Central Queensland	6. Rural 1	QLD	ML308	2310.1	37	0		0		–	
Murrumbidgee	5. Regional 2	NSW	ML116	2215.0	38	74	13	1709	8	–	
Far North Queensland	7. Rural 2	QLD	ML311	2156.6	39	5.2	31	0		–	
Goulburn Valley	5. Regional 2	VIC	ML215	2073.9	40	0		0		740.5	6
Tasmania	5. Regional 2	TAS	ML601	2001.5	41	48.9	24	311.2	32	502.3	11

Gippsland	5. Regional 2	VIC	ML217	1925.5	42	_		0		−	
West Moreton - Oxley	3. Metro 3	QLD	ML305	1896.1	43	0		0		−	
Northern Territory	7. Rural 2	NT	ML701	1886.8	44	0		0		−	
Country South SA	5. Regional 2	SA	ML404	1824.4	45	−		0		−	
New England	5. Regional 2	NSW	ML114	1745.0	46	_		0		−	
Northern Adelaide	3. Metro 3	SA	ML401	1708.8	47	0		0		105.8	36
Macedon Ranges and North Western Melbourne	3. Metro 3	VIC	ML204	1681.5	48	0		350.7	31	−	
South Western Melbourne	2. Metro 2	VIC	ML203	1665.2	49	0		0		−	
South Eastern Melbourne	3. Metro 3	VIC	ML208	1630.5	50	0		200	35	171.4	34
Country North SA	6. Rural 1	SA	ML405	1592.1	51	0		0		−	
Perth South Coastal	4. Regional 1	WA	ML505	1565.9	52	_		471.4	29	−	

Great South Coast	5. Regional 2	VIC	ML212	1316.1	53	0		0		–	
Goldfields - Midwest	7. Rural 2	WA	ML507	1166.8	54	_		0		–	
Southern NSW	5. Regional 2	NSW	ML117	968.8	55	0		0		415.3	18
Kimberley - Pilbara	7. Rural 2	WA	ML508	937.1	56	_		0		0	
Bentley - Armadale	2. Metro 2	WA	ML504	806.0	57	0		11.1	36	69.5	37
Perth North Metro	2. Metro 2	WA	ML502	671.7	58	0		0		57	38
South West WA	5. Regional 2	WA	ML506	485.4	59	–		0		–	
Central and North West Queensland	7. Rural 2	QLD	ML309	382.8	60	_		_		0	
Far West NSW	6. Rural 1	NSW	ML118	–		–		–		0	

Appendix

Appendix Table 3 . Association of population characteristics in Medicare Locals with age-weighted numbers of tests/100,000 persons.

	TTE				TOE				SE				SPECT			
	Beta	**95%CI (Lower-Upper)**		**p**	**Beta**	**95%CI (Lower-Upper)**		**p**	**Beta**	**95%CI (Lower-Upper)**		**p**	**Beta**	**95%CI (Lower-Upper)**		**p**
Women (%)	1.224	1.119	1.340	<0.001	1.119	0.791	1.581	0.525	1.118	0.694	1.799	0.646	1.967	1.598	2.420	<0.001
Non-cardiologists (/1,000)	1.223	1.145	1.306	<0.001	1.199	1.106	1.299	<0.001	1.175	1.032	1.337	0.015	1.021	0.881	1.184	0.781
Older 65years (%)	1.052	1.012	1.093	0.010	0.943	0.870	1.022	0.153	0.989	0.900	1.086	0.817	1.155	1.069	1.247	<0.001
CV disease (%)	0.989	0.962	1.016	0.420	0.936	0.908	0.965	<0.001	0.947	0.903	0.992	0.023	1.024	0.975	1.076	0.335
CV deaths (/1,000)	1.051	0.866	1.274	0.616	0.767	0.578	1.016	0.064	0.977	0.697	1.369	0.892	1.518	1. 009	2.284	0.045
More disadvantaged (%)	0.992	0.985	1.000	0.039	0.982	0.973	0.991	<0.001	0.988	0.976	1.000	0.054	1.008	1.000	1.017	0.059
Peer group																
- Metro 1	Ref	-	-	-	Ref	-	-	-	Ref	-	-	-	Ref	-	-	-
- Metro 2	0.751	0.488	1.156	0.194	0.767	0.530	1.110	0.160	0.689	0.365	1.301	0.251	0.622	0.363	1.066	0.084

Appendix

- Metro 3	0.496	0.304	0.809	0.005	0.443	0.267	0.734	0.002	0.451	0.210	0.969	0.041	0.745	0.390	1.422	0.372
- Regional 1	0.717	0.446	1.150	0.167	0.451	0.284	0.716	0.001	0.310	0.155	0.621	0.001	1.585	0.859	2.925	0.141
- Regional 2	0.552	0.366	0.831	0.004	0.416	0.275	0.627	<0.001	0.571	0.277	1.178	0.129	1.390	0.788	2.450	0.256
- Rural 1	0.529	0.296	0.944	0.031	0.416	0.189	0.916	0.029	†				0.413	0.167	1.022	0.056
- Rural 2	0.270	0.157	0.463	<0.001	0.040	0.016	0.103	<0.001	†				†			
Region																
- Metro	Ref	-	-	-	Ref	-	-	-	Ref	-	-	-	Ref	-	-	-
- Regional	0.800	0.597	1.071	0.133	0.537	0.369	0.782	0.001	0.577	0.350	0.953	0.032	1.888	1. 101	3.240	0.021
- Rural	0.508	0.340	0.758	<0.001	0.287	0.139	0.589	0.001	†				0.265	0.111	0.635	0.003
Cardiologists (/100,000)*					0.338	0.179	0.637	0.001	0.411	0.255	0.661	<0.001				

† Unable to estimate, no testing performed. * Estimates from inflation part of ZINB model.

Appendix

Appendix Table 4 . Characteristics of studies included in the systematic review

Study	Test	AUC Edition	Type of Study	AUC score made by:	Where?	Profession of reviewer	Collection of information for scoring	Methodology used for scoring	Blinded to the Hypothesis of the study	Kappa	Academic or Community setting	Sample	n App tests	n Classified Studies	%Classified Tests	%App of total	%App of classified
Ward	TTE	2007	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	0.82	Academic Medical Center	1553	1228	1385	0.89	0.79	0.89
Willens	TTE	2007	Prospective	Reviewers	Point of service	Echocardiographer	Review of medical records	Single determination	Not specified	Not specified	Both	625	481	526	0.84	0.77	0.91
Dharmarajan	TTE	2007	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Academic Medical Center	58	51	58	1.00	0.88	0.88
Kirkpatrick	TTE	2007	Prospective	Reviewers	Point of service	Echocardiographer	Review of medical records	Single determination	Not specified	0.67	Academic Medical Center	368	206	237	0.64	0.56	0.87
Martin	TTE	2007	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	274	237	268	0.98	0.86	0.88

Appendix

Bhave	TTE	2007	Prospective	Reviewers	Point of service	Sonographers and Physicians	Review of medical records	Not specified	Not specified	0.84	Academic Medical Center	258	199	221	0.86	0.77	0.90
Rao	TTE	2007	Prospective	Not specified	Point of care	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Community Setting	772	533	716	0.93	0.69	0.74
Aggarwal	TTE	2007	Retrospective	reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	0.55	Academic Medical Center	329	278	299	0.91	0.84	0.93
Gathak	TTE	2007	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Not specified	Not specified	Not specified	Academic Medical Center	431	364	394	0.91	0.84	0.92
Rahimi1	TTE	2007	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	177	143	164	0.93	0.81	0.87
Parikh	TTE	2007	Retrospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	384	333	336	0.88	0.87	0.99
Bhatia	TTE	2007	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassif	Not specified	Not spe	Both	450	288	347	0.77	0.64	0.83

Appendix

								ied/disagreement		cified							
Alqarqaz	TTE	2007	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	170	131	147	0.86	0.77	0.89
Silverman	TTE	2007	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Academic Medical Center	485	442	485	1.00	0.91	0.91
Bailey	TTE	2007	Retrospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	1080	933	945	0.88	0.86	0.99
Willens	TTE	2011	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Both	625	479	617	0.99	0.77	0.78
Patil	TTE	2011	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	1820	1493	1812	1.00	0.82	0.82
Ballo	TTE	2011	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Multiple determination	Not specified	0.83	Community Setting	931	739	920	0.99	0.79	0.80

Appendix

Mansour	TTE	2011	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	1553	1253	1525	0.98	0.81	0.82
Matulevicius	TTE	2011	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	0.8	Academic Medical Center	535	491	535	1.00	0.92	0.92
Bhatia2	TTE	2011	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	1318	1105	1312	1.00	0.84	0.84
Rao	TEE	2007	Retrospective	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Academic Medical Center	1235	1156	1235	1.00	0.94	0.94
Ogbara	TEE	2007	Prospective	Reviewers	Not specified	Not specified	Review of medical records	Not specified	Not specified	Not specified	Academic Medical Center	389	321	389	1.00	0.83	0.83
Bhatia	TEE	2007	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	202	156	166	0.82	0.77	0.94
Grewal	TEE	2011	Retrospective	Reviewers	Point of service	Not specified	Review of medical records	Not specified	Not specified	Not spe	Academic Medical Center	671	639	659	0.98	0.95	0.97

Appendix

										cified							
McCully	SE	2008	Retrospective	Reviewers	Point of service	Nurses	Review of medical records	Review only for unclassified/disagreement	Not specified	0.72	Academic Medical Center	298	159	241	0.81	0.53	0.66
Mansour	SE	2008	Retrospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	289	180	253	0.88	0.62	0.71
Bhatia	SE	2008	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	252	104	126	0.50	0.41	0.83
Willens	SE	2008	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Both	209	104	189	0.90	0.50	0.55
Lin	SE	2008	Prospective	Ordering physician	Point of care	Physicians	criteria entered by ordering physician	Single determination	Not specified	Not specified	Community Setting	111	50	92	0.83	0.45	0.54
Schmitz	SE	2008	Retrospective	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Academic Medical Center	300	194	226	0.75	0.65	0.86

Appendix

Cortigiani	SE	2011	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Community Setting	1552	984	1552	1.00	0.63	0.63
Bhattacharyya	SE	2011	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	both	100	49	100	1.00	0.49	0.49
Gibbons1	SPECT	2005	Retrospective	reviewers	Point of service	Nurses	Review of medical records	Review only for unclassified/disagreement	Not specified	0.56	Academic Medical Center	284	182	253	0.89	0.64	0.72
Mehta	SPECT	2005	Retrospective	Reviewers	Not specified	Not specified	Review of medical records	Single determination	Not specified	Not specified	Academic Medical Center	1209	940	1173	0.97	0.78	0.80
Hendel	SPECT	2005	Prospective	Reviewers	Point of service	Nurses	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Community Setting	6351	4192	5906	0.93	0.66	0.71
Gibbons2	SPECT	2005	Retrospective	Reviewers	Point of service	Nurses	Review of medical records	Review only for unclassified/disagreement	Not specified	0.68	Academic Medical Center	284	188	241	0.85	0.66	0.78
Carryer	SPECT	2005	Retrospective	reviewers	Point of service	Nurses	Review of medical records	Review only for unclassif	Not specified	Not spe	Academic Medical Center	281	179	250	0.89	0.64	0.72

Appendix

								ied/disagreement		cified							
Gupta	SPECT	2005	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Not specified	Not specified	Not specified	Academic Medical Center	314	263	314	1.00	0.84	0.84
Gibbons3	SPECT	2005	Prospective	Reviewers	Point of service	Nurses	Review of medical records	Review only for unclassified/disagreement	Not specified	0.74	Academic Medical Center	273	164	232	0.85	0.60	0.71
Gholamrezanezhad	SPECT	2005	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Multiple determination	Not specified	0.63	Both	291	211	279	0.96	0.73	0.76
Druz	SPECT	2005	Prospective	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Academic Medical Center	585	370	570	0.97	0.63	0.65
Soine1	SPECT	2005	Retrospective	Reviewers	Point of service	Not specified	Review of medical records	Not specified	Not specified	Not specified	Academic Medical Center	1377	950	1377	1.00	0.69	0.69
Koh	SPECT	2009	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	0.64	Academic Medical Center	1623	1331	1574	0.97	0.82	0.85

Appendix

Nelson1	SPECT	2009	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Community Setting	150	101	148	0.99	0.67	0.68
Koh	SPECT	2009	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	176	106	176	1.00	0.60	0.60
Winchester	SPECT	2009	Retrospective	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Not specified	Community Setting	332	259	328	0.99	0.78	0.79
Doukky	SPECT	2009	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Single determination	Not specified	0.83	Community Setting	1511	779	1491	0.99	0.52	0.52
Moralidis	SPECT	2009	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	3032	2208	3008	0.99	0.73	0.73
Aldweib	SPECT	2009	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Academic Medical Center	1199	740	1194	1.00	0.62	0.62
Ayyad 1	CCT	2006	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Both	763	530	715	0.94	0.69	0.74

Appendix

Miller	CCT	2006	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	0.31	Academic Medical Center	251	69	136	0.54	0.27	0.51
Murphy	CCT	2006	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	267	126	189	0.71	0.47	0.67
El Sibai	CCT	2006	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Academic Medical Center	100	8	100	1.00	0.08	0.08
Chinnaiyan	CCT	2006	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Single determination	Not specified	Not specified	Both	25387	5053	12853	0.51	0.20	0.39
Rich	CCT	2006	Prospective	Reviewers	Point of service	Not specified	Review of medical records	Not specified	Not specified	Not specified	Academic Medical Center	1216	503	1069	0.88	0.41	0.47
Mazimba	CCT	2006	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Not specified	Not specified	0.84	Academic Medical Center	243	36	243	1.00	0.15	0.15
Wasfy	CCT	2006	Prospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	Not specified	Academic Medical Center	267	119	189	0.71	0.45	0.63

Appendix

Cullen	CCT	2010	Retrospective	Reviewers	Point of service	Physicians	Review of medical records	Review only for unclassified/disagreement	Not specified	0.55	Academic Medical Center	251	85	212	0.84	0.34	0.40

Appendix

Appendix Table 5 . Clinical characteristics of patients according to appropriate use category.

	No TTE	**Inappropriate**	**Uncertain**	**Appropriate**	**p[(a)]**	**p[(b)]**	**p[(c)]**	**p[(d)]**	**p[(e)]**
n	409	41	19	81					
Age (median [IQR])	79 [70, 84]	69 [63, 76]	76 [69, 79]	75 [68, 80]	**<0.01**	0.06	0.14	**0.02**	0.82
Male sex, n (%)	222 (54.3)	25 (61.0)	9 (47.4)	51 (63.0)	0.38	0.46	0.48	0.99	0.32
Body mass index, median [IQR]	26 [23, 30]	27 [23, 33]	28 [24, 34]	28 [25, 31]	0.06	0.73	0.64	0.44	0.77
Systolic blood pressure in mmHG, median [IQR]	120 [105, 135]	120 [110, 130]	128 [106, 140]	121 [110, 136]	0.64	0.67	0.57	0.39	0.84
Creatinine levels in µ/L, median [IQR]	112 [86, 148]	88 [72, 117]	85 [67, 111]	106 [86, 126]	**<0.01**	0.06	0.81	**0.04**	0.08
Left ventricular ejection fraction %, median [IQR]	0.45 [0.32, 0.60]	0.39 [0.35, 0.50]	0.43 [0.25, 0.51]	0.46 [0.35, 0.60]	0.14	0.06	0.89	**0.03**	0.11
MAGGIC score, median [IQR])	28 [23, 33]	24 [19, 27]	27 [21, 29]	25 [20, 28]	**<0.01**	0.48	0.26	0.39	0.52

Risk of death 1 year, median [IQR]	0.21 [0.13, 0.32]	0.15 [0.09, 0.19]	0.19 [0.11, 0.24]	0.16 [0.10, 0.21]	<**0.01**	0.48	0.25	0.41	0.52
Risk of death 3 years, median [IQR]	0.46 [0.32, 0.63]	0.34 [0.23, 0.43]	0.43 [0.28, 0.50]	0.37 [0.25, 0.46]	<**0.01**	0.48	0.26	0.39	0.52
Betablockers, n (%)	225 (55.0)	26 (63.4)	12 (63.2)	52 (64.2)	0.34	0.99	1.00	1.00	1.00
Diuretics, n (%)	369 (90.2)	38 (92.7)	15 (78.9)	69 (85.2)	0.22	0.30	0.27	0.37	0.75
Mineralocorticoids, n (%)	124 (30.3)	19 (46.3)	5 (26.3)	25 (30.9)	0.20	0.17	0.23	0.14	0.91
Angiotensin-converting-enzyme inhibitor/ angiotensin-II receptor blockers, n (%)	304 (74.3)	34 (82.9)	15 (78.9)	57 (70.4)	0.48	0.29	0.99	0.20	0.64
Calcium antagonists, n (%)	82 (20.0)	5 (12.2)	3 (15.8)	16 (19.8)	0.70	0.63	0.70	0.43	1.00
Antiarrhythmics, n (%)	44 (10.8)	2 (4.9)	2 (10.5)	13 (16.0)	0.31	0.19	0.59	0.09	0.73
Digoxin, n (%)	84 (20.5)	8 (19.5)	6 (31.6)	13 (16.0)	0.49	0.30	0.48	0.82	0.22
Statins, n (%)	210 (51.3)	16 (39.0)	10 (52.6)	48 (59.3)	0.21	0.11	0.48	0.06	0.79

Appendix

Hypertension, n (%)	292 (71.4)	24 (58.5)	11 (57.9)	66 (81.5)	0.03	**0.01**	1.00	**0.01**	0.06
Dyslipidaemia, n (%)	194 (47.4)	14 (34.1)	11 (57.9)	52 (64.2)	**0.01**	**0.01**	0.15	**<0.01**	0.80
History of angina, n (%)	119 (29.1)	8 (19.5)	9 (47.4)	38 (46.9)	**<0.01**	**0.01**	0.06	**0.01**	1.00
History of atrial fibrillation, n (%)	203 (49.6)	19 (46.3)	9 (47.4)	47 (58.0)	0.51	0.41	1.00	0.30	0.56
Arrhythmia, n (%)	68 (16.6)	5 (12.2)	5 (26.3)	16 (19.8)	0.51	0.38	0.32	0.43	0.75
Cardiomyopathy, n (%)	153 (37.4)	10 (24.4)	9 (47.4)	29 (35.8)	0.29	0.19	0.14	0.28	0.50
Dilated cardiomyopathy, n (%)	115 (28.1)	8 (19.5)	8 (42.1)	22 (27.2)	0.34	0.19	0.13	0.48	0.32
Hypertrophic cardiomyopathy, n (%)	25 (6.1)	1 (2.4)	1 (5.3)	6 (7.4)	0.77	0.60	0.54	0.42	1.00
Restrictive cardiomyopathy, n (%)	13 (3.2)	1 (2.4)	0 (0.0)	1 (1.2)	0.89	1.00	1.00	1.00	1.00
Deep vein thrombosis, n (%)	20 (4.9)	1 (2.4)	0 (0.0)	4 (4.9)	0.95	0.84	1.00	0.66	1.00
Angioplasty, n (%)	78 (19.1)	9 (22.0)	5 (26.3)	19 (23.5)	0.71	0.93	0.97	1.00	1.00

Appendix

Myocardial infarction, n (%)	178 (43.5)	11 (26.8)	11 (57.9)	30 (37.0)	0.07	0.07	**0.04**	0.36	0.16
Cerebrovascular disease, n (%)	93 (22.7)	6 (14.6)	0 (0.0)	25 (30.9)	**0.01**	**<0.01**	0.16	0.09	**<0.01**
Renal disease, n (%)	106 (25.9)	5 (12.2)	5 (26.3)	21 (25.9)	0.28	0.20	0.32	0.13	1.00
Valvular disease, n (%)	152 (37.2)	15 (36.6)	7 (36.8)	29 (35.8)	1.00	0.99	1.00	1.00	1.00
Cardiac catheterisation, n (%)	53 (13.0)	2 (4.9)	2 (10.5)	15 (18.5)	0.20	0.09	0.59	0.05	0.52
Diabetes mellitus, n (%)	166 (40.6)	12 (29.3)	9 (47.4)	32 (39.5)	0.48	0.35	0.28	0.36	0.71
Chronic obstructive pulmonary disease, n (%)	167 (40.8)	21 (51.2)	6 (31.6)	34 (42.0)	0.49	0.34	0.25	0.44	0.57
NYHA class, n (%)					**0.01**	0.63	0.46	0.73	0.39
1	78 (19.1)	10 (24.4)	2 (10.5)	24 (29.6)					
2	111 (27.1)	12 (29.3)	9 (47.4)	28 (34.6)					
3	142 (34.7)	16 (39.0)	7 (36.8)	25 (30.9)					
4	78 (19.1)	3 (7.3)	1 (5.3)	4 (4.9)					

Appendix

NHYA class 3 and 4, n (%)	220 (53.8)	19 (46.3)	8 (42.1)	29 (35.8)	**0.02**	0.52	0.98	0.35	0.80
Active change in management after follow-up TTE (%)	_	10 (24.4)	2 (10.5)	18 (22.2)	_	0.45	0.37	0.97	0.41

Data in bold denote statistically significant results

Appendix

Appendix 6- 1 . Interview Schedule

Opening

1. (Establish rapport) Hello Dr_____, I am Ricardo Fonseca, the PhD candidate who contacted you for the appropriate use of echocardiography study. How are you?
2. (Purpose) I would like to ask you some questions about your background, some experiences you have had when requesting an echocardiogram and some of your possible decisions in some situations when a patient was referred for an echocardiogram.
3. (Motivation) Your comments will help us to learn more about factors that affect the process of decision-making and appropriate use of echocardiography, and to develop more efficient recommendations or educational campaigns to avoid inappropriate testing.
4. (Time line) The interview should take about 30 minutes. Are you available to respond to some questions at this time?

(Transition: Let me begin by asking you some questions about your background and experiences with patients who needed echocardiograms)

Body

1. (Topic) Background
    a. What is your specialty?
    b. Do you usually have patients who need echocardiograms? Or is it rare?
    c. During a month, how many echocardiograms you usually request? No need for an exact number, just an approximation.
    d. What is more common to you: requesting an echocardiogram for inpatients or outpatients? For new findings or follow-up?

2. (Topic) Experiences when requesting an echocardiogram:
    a. In what percentage of cases, you order a test due to a cause other than medical indication?
    b. There is an interesting manuscript written more than 30 years ago in the NEJM by an author called Joseph Hardison. I have provided the 10 questions he asked his colleagues. Could you please indicate if you have experienced the following

reasons to perform an echocardiogram? Can you rank these in the order of frequency or the percentage of requests in each category?

"To be complete"	-
A superior or colleague said the patient needed it.	-
"Fear" of getting into trouble (perhaps by missing something)	-
Because the patient is here and it is better to order everything at once	-
Because the patient is hospitalised	-
Academic reasons	-
Protocols	-
Malpractice	-
"How do I know the patient does have it?"	-
"If the patient was my father/mother/partner/relative, I would want it done"	-

- Do you have any other reason to order an echocardiogram?

3. (Topic) I would now like to provide you with some clinical situations and discuss whether or not you would order an echocardiogram in these circumstances and the reasons why. Can you please talk through you're decision making/reasoning process when considering each of the five scenarios presented here. (The following scenarios will be presented in writing to each participant).

a. 44 year old business man who presented to GP for health insurance check-up and was found to have an incidental murmur. Patient is slightly overweight. Rest of physical exam is normal. He has reasonable exercise tolerance and walks his dog 5 times a week without any problems. His echocardiogram shows mild mitral valve prolapse and mild mitral regurgitation. You see him in one year time for follow-up. He is doing well; continues to walk his dog without any problems. However has gained 5

kilograms of weight. His physical examination is normal except for the cardiac murmur found previously.

b. 37 year old male with high grade fevers, IVDU. Cellulitis around injection site. Examination shows low body weight. Normal cardiovascular exam. Blood cultures negative so far. Query echocardiogram for infective endocarditis?
c. 50 year old female with dilated cardiomyopathy well treated with medications. Ejection fraction 40%. Mildly dilated left ventricle. Exercise tolerance 1 km on flat. Mild shortness of breath up hills, similar to 6 months ago. Presents for follow-up after 6 months for titration of medications.
d. 88 year old female, living in low level care hostel, no significant co-morbidities, and walks assisted with walk-stick. Non-severe aortic stenosis on echocardiogram 14 months ago. Reports she is asymptomatic for annual follow –up and echocardiogram.
e. 20 year old male, presents with chest pain following recent flu. ECG shows widespread ST elevation consistent with pericarditis. WCC, CRP, ESR, are all elevated. Echocardiogram is normal. Except for small pericardial effusion. Patient is followed up in cardiology clinic three months later following a course of NSAIDs and colchicine. He feels well, has not had chest pain for two months. Repeat echocardiogram to assess resolution for pericardial effusion. .

For each scenario provide prompts such as:

What is the most important consideration in making this decision (i.e. age, severity of symptoms, previous test results)? What other factors would influence your decision to order an echocardiogram?

If an echocardiogram is recommended by the interviewee, how urgently does the test need to be conducted?

4. (Topic) What do you do if the echocardiogram result shows an incidental finding? (Example scenario: 23 year old female, G1 P0, 27 weeks pregnant, seen in obstetric clinic with incidental murmur. Referred for echo which shows query small linear aortic valve mass (this could be an artefact).)

5. (Topic) Assessing awareness of guidelines and AUC: When making decisions about ordering an echocardiogram, do you consider any guidelines/criteria/protocols? Are you familiar with the Guidelines for Clinical Application of Echocardiography? And the Appropriateness Criteria for Echocardiography?

6. (Topic) Patient involvement in decision making: When making the decision to order an echocardiogram do you discuss the possibility of requesting an echo with your patient? Does this discussion involve consideration of possible benefits and harms? If the response is no to patient involvement / discussion of benefits-harms then follow up with: under what circumstances would you involve the patient in the decision making process about their care?

7. (Topic) Patient desires: Have you ever had a patient ask for an echocardiogram? If yes, then ask: How does this request impact your decision making about ordering an echocardiogram?

8. (Topic) Harms of echocardiography: Are you aware of harms of echocardiogram? Could you say some?

9. (Topic) Awareness of costs: Do you know how much Medicare pays for a transthoracic echocardiogram? Trans-oesophageal? Stress echo? Other cardiac imaging (SPECT, cardiac tomography-angiography)? Do you take this into account when you are ordering a test?

10. (Topic) When you order an echocardiogram, are you always aware of its result? Where are the test results recorded? How do the test results impact on your management of the patient?

(Transition: Well, it has been a pleasure finding out more about your thoughts when you are considering the possibility of requesting an echocardiogram)

Closing

- (Maintain Rapport) I appreciate the time you took for this interview. Is there anything we haven't discussed that you consider influences your decision making about ordering echocardiograms?
- (Action to be taken) I should have all the information I need. Would it be alright to contact you again if I have any more questions? Thanks again. I look forward to analysing the data and to finding efficient recommendations to improve appropriate use criteria for testing.

Appendix 6- 2 . Clinical scenarios

a. 44 year old business man who presented to GP for health insurance check-up and was found to have an incidental murmur. Patient is slightly overweight. Rest of physical exam is normal. He has reasonable exercise tolerance and walks his dog 5 times a week without any problems. His echo shows mild mitral valve prolapse and mild mitral regurgitation. You see him in one year time for follow-up. He is doing well; continues to walk his dog without any problems. However has gained 5 kilograms of weight. His physical examination is normal except for the cardiac murmur found previously.
b. 37 year old male with high grade fevers, IVDU. Cellulitis around injection site. Examination shows low body weight. Normal cardiovascular exam. Blood cultures negative so far. Query echo for infective endocarditis?
c. 50 year old female with dilated cardiomyopathy well treated with medications. Ejection fraction 40%. Mildly dilated left ventricle. Exercise tolerance 1 km on flat. Mild shortness of breath up hills, similar to 6 months ago. Presents for follow-up after 6 months for titration of medications.
d. 88 year old female, living in low level care hostel, no significant co-morbidities, and walks assisted with walk-stick. Non-severe aortic stenosis on echo 14 months ago. Reports she is asymptomatic for annual follow –up and echo.
e. 20 year old male, presents with chest pain following recent flu. ECG shows widespread ST elevation consistent with pericarditis. WCC, CRP, ESR, are all elevated. Echocardiogram is normal. Except for small pericardial effusion. Patient is followed up in cardiology clinic three months later following a course of NSAIDs and colchicine. He feels well, has not had chest pain for two months. Repeat echo to assess resolution for pericardial effusion. .

For each scenario provide prompts such as:

What is the most important consideration in making this decision (i.e. age, severity of symptoms, previous test results)? What other factors would influence your decision to order an echo?

If an echo is recommended by the interviewee, how urgently does the test need to be conducted?

Appendix 6- 3 . Approach letter

Dear Doctor _____.

Researchers at the Menzies Institute for Medical Research, University of Tasmania, are investigating the factors leading to a physician requesting echocardiograms and its relationship with the appropriateness of the study. The aim of this work is to understand the decision making process around cardiac imaging. You are being invited to participate in this important research because you have:

- Requested an echocardiogram in the last year at the Royal Hobart Hospital.

Health care expenditure has increased considerably, jeopardising the limited healthcare resources. The appropriate use of medical services has been one of the major concerns to secure a sustainable health care system. In the context of limited resources, health resources should be used appropriately. Although guidelines for the use of cardiac imaging have been developed and the appropriate use criteria have been used for a decade, around 10-15%% of the echocardiograms are still inappropriate. In Australia, up to 20% of transthoracic echocardiograms are inappropriate. The Medicare reimbursement for all echocardiography modalities in 2014 in Australia was over 235 million dollars, and around 178 million dollars were reimbursed only for transthoracic echocardiography (TTE): improving appropriate TTE ordering may result in significant health care savings.

The researchers want to talk to doctors who have requested echocardiograms at the Royal Hobart Hospital. The research will be conducted using face to face interviews that will take place at a central location near you and will last for 30 minutes. Understanding the factors that influence decisions about cardiac imaging will contribute to the development of more efficient recommendations or educational campaigns to avoid inappropriate testing.

Doctor _____, being involved in this research will provide you with an opportunity to make a difference to our health system and contribute to how future research in this area will be done.

Participation in this study is voluntary. It is completely up to you whether or not you participate. All data will be kept confidential and access limited to the researchers. Doctors will not be identifiable in any reporting or publication.

Appendix

If you would like to participate in the interviews, please read the information sheet attached and contact Doctor Ricardo Fonseca at Ricardo.Fonseca@utas.edu.au.

Kind regards,

Dr Tom Marwick

MBBS, PhD, MPH

Adjunct Professor, Menzies Institute for Medical Research

Appendix 6- 4 . Information sheet

Understanding Cardiac imaging decision making: appropriate testing determinants

Information sheet for health care professionals individual interviews

What is the purpose of this study?

The purpose of this study is to understand the factors that influence physicians in their decision-making related to echocardiography and its impact on the appropriateness of the test.

Aim

Our research aim is to explore the various influences on appropriate cardiac imaging decision-making by doctors to identify areas in which interventions are needed to improve health outcomes and health care expenditure, and areas in which further research is needed.

Why have I been invited to participate?

You have been invited to participate in this study because you:

- Have referred a patient for echocardiography in the past year.
- You are a physician currently working at the Royal Hobart Hospital.

If you choose to participate, this decision, along with personal information that you provide the researchers, will be kept confidential. Participation in this study is voluntary. It is completely up to you whether or not you participate.

What will I be asked to do?

We will ask you to complete an interview. The interview may be conducted face to face. The interview may take up to 30 minutes and will be recorded. The interview will be about requesting cardiac imaging and its appropriateness and will cover topics such as:

- Knowledge and understanding about appropriateness criteria of cardiac imaging.
- Decision making about echocardiography.
- Factors that influence cardiac imaging testing practices.
- Possible solutions to some of the challenges in referral for echocardiography.
- Decisions after an inappropriate test which shows abnormal results.

Are there any possible benefits from participation in this study?

- The identification of common challenges when requesting a test, will help us to understand why doctors request tests and its impact on appropriateness. The results may also influence policy and clinical practice interstate and internationally.
- It is an opportunity to provide input into future research that will be done in this area.

Are there any possible risks from participation in this study?

There are no foreseeable risks associated with your participation in this study.

What if I change my mind during or after the study?

You are free to withdraw from the study at any time, and you can do so without providing an explanation. If you choose to withdraw after you have participated in the interview we will delete your interview data.

What will happen to the information when this study is over?

The study data will be kept for 5 years from the first date of publication in a secure location in the Menzies Institute for Medical Research at the University of Tasmania in Hobart. The researchers will have the right to access this data for the purpose of publication. Your data will be treated in a confidential manner. The anonymised transcripts of the interview may be passed to other researchers for use in other projects but before doing so, we will endeavour to remove all identifying information.

How will the results of the study be published?

The results of this study will be published:

- As a summary on the Menzies Institute for Medical Research website http://www.menzies.utas.edu.au.
- In peer reviewed scientific journals

Participants or their practices will not be identifiable in the publication of the results.

What if I have questions about this study?

If you have any questions about this study or if you would like to contribute to this area of research in other ways, you can contact:

- Dr Ricardo Fonseca (PhD candidate, Menzies Institute for Medical Research), email: Ricardo.Fonseca@utas.edu.au

- Dr. Kim Jose (Post Doctoral Research Fellow, Menzies Institute for Medical Research)
  Email: kim.jose@utas.edu.au
- Professor Tom Marwick (Menzies Institute for Medical Research ,University of Tasmania)
  Email: tom.marwick@utas.edu.au

'This study has been approved by the Tasmanian Health and Medical Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study should contact the Executive Officer of the HREC (Tasmania) Network on (03) 6226 7479 or email human.ethics@utas.edu.au. The Executive Officer is the person nominated to receive complaints from research participants. You will need to quote HREC project number H0015516.

This information sheet is for you to keep. You will be provided with a written consent form, which you will need to complete and sign prior to participation in this study.

Appendix 6- 5 . Consent form

Understanding Cardiac imaging decision making: appropriate testing determinants

Consent form for individual interviews

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves engaging in an interview. The interview will be face to face. Interviews will be recorded.
5. There are no foreseeable risks associated with my participation in this study. The information will be stored in de-identified format and information that permits re-identification will be kept securely.
6. I understand that all research data will be securely stored on the Menzies Institute for Medical Research, University of Tasmania premises for fifteen years from the publication of the study results, and will then be destroyed.
7. Any questions that I have asked have been answered to my satisfaction.
8. I understand that the researcher(s) will maintain confidentiality and that any information I supply to the researcher(s) will be used only for the purposes of the research.
9. I understand that the results of the study will be published in such a way that I cannot be identified as a participant. My practice will also not be identifiable.
10. I understand that my participation is voluntary and that I may withdraw at any time without any effect.

Participant's name: _____

Participant's signature: _____Date: _____

Statement by Investigator

☐ I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

☐ The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name: _____

Investigator's signature: _____Date: _____

Appendix Table 6. Diagnostic tests for each of the questions in each of the groups

		Sensitivity Estimate[lower-upper]	Specificity Estimate[lower-upper]	Odds ratio Estimate [lower-upper]	Positive predictive value Estimate [lower-upper]	Negative predictive value Estimate [lower-upper]	Positive likelihood ratio Estimate[lower-upper]	Negative likelihood ratio Estimate[lower-upper]
Prospective cohort	≥2 affirmative answers	0.84[0.68,0.94]	0.87[0.83,0.90]	33.96[13.61,84.78]	0.34[0.24,0.44]	0.98[0.97,1.00]	6.34[4.83,8.34]	0.19[0.09,0.39]
	Question 1 answered as "yes"	0.96[0.82,0.99]	0.79[0.75,0.83]	66.72[15.77,282.32]	0.27[0.19,0.35]	0.99[0.98,1.00]	4.55[3.75,5.53]	0.07[0.02,0.26]
	Question 2 answered as "yes"	0.73[0.39,0.94]	0.89[0.82,0.94]	21.33[5.10,89.27]	0.35[0.16,0.57]	0.98[0.93,0.99]	6.54[3.60,11.91]	0.31[0.12,0.81]
	Question 3 answered as "yes"	0.68[0.50,0.82]	0.79[0.75,0.82]	7.74[3.75,15.95]	0.20[0.14,0.29]	0.97[0.95,0.98]	3.19[2.40,4.23]	0.41[0.26,0.66]
	Question 4 answered as "yes"	0.11[0.03,0.25]	1.00[0.99,1.00]	Infinite	1.00[0.28,1.00]	0.93[0.91,0.95]	Infinite	0.89[0.80,1.00]
External retrospective validation cohort	≥2 affirmative answers	0.80[0.69,0.88]	0.95[0.84,0.97]	83.3[43.13,159.83]	0.62[0.51,0.72]	0.98[0.97,0.99]	17.41[12.46,24.32]	0.21[0.13,0.33]
	Question 1 answered as "yes"	0.95[0.87,0.99]	0.91[0.89,0.93]	177.99[63.19,501.32]	0.49[0.41,0.58]	0.99[0.99,1.00]	10.44[8.33,13.08]	0.06[0.02,0.15]
	Question 2 answered as "yes"	0.53[0.41,0.65]	0.96[0.94,0.97]	26.74[15.09,47.37]	0.55[0.43,0.66]	0.96[0.94,0.97]	13.01[8.76,19.32]	0.49[0.38,0.62]
	Question 3 answered as "yes"	0.37[0.28,0.51]	0.93[0.91]0.94]	7.97[4.67,13.61]	0.33[0.23,0.44]	0.94[0.92,0.96]	5.28[3.62,7.69]	0.66[0.55,0.79]
	Question 4 answered as "yes"	0.23[0.14,0.38]	1.00[0.99,1.00]	Infinite	1.00[0.73,1.00]	0.93[0.91,0.95]	Infinite	0.77[0.68,0.87]